



Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Méréstechnika és Információs Rendszerek Tanszék

Kaotikus rendszerek zajos parciális
megfigyeléseiből való előrejelzés empirikus módú
dekompozíció alkalmazásával

TDK dolgozat

Készítette:

Lichter Bertalan

Konzulens:

Hadházi Dániel

2024

Tartalomjegyzék

1. Bevezetés és háttérismeretek	1
1.1. Nemlineáris dinamika és káosz	1
1.1.1. Az attraktor	2
2. Előrejelzés parciális megfigyelésekből	3
2.1. Attraktor-rekonstrukció és Takens tétele	3
2.2. Kaotikus rendszerek előrejelzése	5
2.2.1. Next Generation Reservoir Computing (NG-RC) [7]	5
2.2.2. NG-RC-Chebyshev modell (NG-RC-Ch) [17]	6
2.3. Előrejelzés zajos megfigyelésekből	7
3. Empirikus módú dekompozíció (EMD)	9
3.1. Az EMD algoritmus	9
3.2. Együttes Empirikus Módú Dekompozíció (EEMD) [22]	10
3.3. Teljes EEMD (CEEMD) [23]	10
3.4. Adaptív zajos CEEMD (CEEMDAN) [20]	10
3.5. Továbbfejlesztett CEEMDAN (ICEEMDAN) [6]	11
4. Radiális Bázisfüggvény Hálózatok	14
4.1. Radiális Bázisfüggvény Hálózat (RBFN)	14
4.1.1. K-Means klaszterező algoritmus [13]	14
4.2. Autóregresszív RBFN (RBF-AR) [12]	15
4.3. Fuzzy RBF-AR (FRBF-AR) [10]	15
5. Az ICEEMDAN-FRBF-AR modell	17
5.1. Zajcsökkentés	17
5.2. Tanulás és előrejelzés	18
5.2.1. Az F-RBF-AR modell Takens-beágyazás mellett	20
5.2.2. Intuíció többféleképpen	21
5.2.3. NG-RC kiterjesztés	22
5.2.4. Hiperparaméterek és megválasztásuk	22
5.2.4.1. Zajcsökkentéskor elhagyott IMF-ek száma	23
5.2.4.2. RBF középpontok száma	23
5.2.4.3. Visszatekintési ablak az AR folyamatoknál (m)	23
6. Eredmények	24
6.1. A vizsgált adathalmaz	24
6.2. Előrejelzések pontosságainak számszerűsítése	24
6.2.1. Determinációs együttható	24
6.2.2. Átlagos Arkusztangens Abszolút Százalékos Hiba (MAAPE) [11]	24
6.2.3. Normalizált Gyök Átlagos Négyzetes Eltérés	25

6.3. Az ICEEMDAN-FRBF-AR modell előrejelzése	25
7. Összefoglalás	26

Kivonat

A minket körülvevő világban sok nemlineáris rendszer által generált adatsorral találkozhatunk, amelyek között találhatók kaotikusak is, mivel a természetben a nemlinearitás és a kaotikusság a jellemző. Ez releváns, többek közt az időjárási megfigyeléseknél, agyi jelfeldolgozásban [12] vagy akár pénzügyi idősoroknál [21], illetve a villamosmérnöki tudományokban [18]. Sokszor azonban ezeknek a rendszereknek nem tudjuk megfigyelni minden változóját, így csak egy-egy parciális megfigyelés áll rendelkezésünkre.

Erre a problémára több, gépi tanuláson alapuló módszert javasoltak az utóbbi időkben, ilyen az NGRC [7] (next-generation reservoir computing), illetve az AR-RBFN [15] (autoregressive radial basis-function network). Ezekkel a modellekkel előre lehet jelezni teljes attraktorokat, vagy akár parciális megfigyeléseket olyan esetekben, amikor zajmentes a megfigyelés. A valódi problémáknál azonban elkerülhetetlen megfigyelési zaj jelenlétével drasztikusan csökken a predikciós pontosságuk. Ebben a dolgozatban egy lehetséges megoldást javasolunk a problémára. Ennek lényege, hogy empirikus módú dekompozícióval [9] (EMD) a megfigyeléseket olyan idősorokra dekomponáljuk, melyekkel leegyszerűsítjük a problémát, és szeparáljuk a megfigyelési zajt, így lehetővé téve ezeknek külön-külön vett előrejelezhetőségét. Kihasználva azt, hogy az EMD által felbontott alapvető módusú függvények (intrinsic mode function - IMF-ek) szuperpozíciója megegyezik a felbontott jellel, az egyes IMF-ek előrejelzéseit összeadva megkapjuk a végleges előrejelzést. A kidolgozott megoldás viselkedését a Lorenz rendszer idősorainak predikciójával vizsgáljuk, mivel megfelelő paraméterek megválasztása mellett ez bizonyítottan kaotikus, továbbá fizikai, folyadékdinamikai rendszert modellez. A parciális megfigyelés ez esetben a Lorenz rendszer x változója. Analizáljuk az egyes modellek előrejelzéseinek pontosságát zajmentes, illetve zajos környezetben. Arra a következtetésre jutunk, hogy rövidtávon az NG-RC alapú modellek a legpontosabbak zajmentes és zajos környezetben is, azonban hosszabb távon, zajos környezetben, a javasolt ICEEMDAN-FRBF-AR modell jobb pontosságot ér el.

1. fejezet

Bevezetés és háttérismeretek

1.1. Nemlineáris dinamika és káosz

Többféle differenciálegyenletet ismerünk, amelyeket kategorizálhatunk többek között változók száma, rendjük, linearitásuk vagy polinomiális esetben fokuk szerint. Ebben a dolgozatban **nemlineáris differenciálegyenletekkel** foglalkozunk, amelyek rendelkeznek olyan változóval, melyre igaz, hogy deriváltja nem írható fel a többi változó értékének és egy konstansnak a lineáris kombinációjaként. Egy egyszerű esetben ez lehet egy változó, amely deriváltja két változó szorzatával arányos. Ilyen például az alábbi differenciálegyenletrendszer:

$$\begin{cases} \dot{x} = 2x \cdot y \\ \dot{y} = 1 + x \end{cases}$$

A fizikában számos helyen találkozhatunk nemlineáris rendszerekkel: ilyenek például a Navier-Stokes-egyenletek¹, amik a folyadék-dinamika törvényeit írják le, de ilyen a Lorenz-rendszer is, amely az atmoszférikus konvekciót hivatott modellezni.

$$\text{Lorenz rendszer} = \begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(\rho - z) - y \\ \dot{z} = xy - \beta z \end{cases}$$

Utóbbi rendszer igen híressé vált egy különlegesebb tulajdonsága révén: bizonyos paraméterek² és kezdeti feltételek mellett kaotikus viselkedést mutat. Érdekes, hogy bár igen szélesen vizsgált és ismert jelenség, a *káosznak* nincs igazán egységes definíciója; legtöbbször csak úgy definiálják, mint olyan nemlineáris dinamikus rendszerek, melyekre az alábbi tulajdonságok teljesülnek: topológiailag tranzitív, rendelkezik pontok sűrű halmazával, melyek periodikus pályákat alkotnak, illetve érzékeny a kezdeti feltételekre (utóbbi pillangó-effektusként is ismert). Fontos kiemelni, hogy a közhiedelemmel ellentétben ezek teljességgel determinisztikus rendszerek, és nem jelenik meg bennük semmilyen sztochasztikus elem.

A kaotikus rendszereknek egy fontos tulajdonsága az ún. *Ljapunov-exponens* (λ), amely jellemzi az egymáshoz infinitezimálisan közeli kezdeti pontokból induló trajektóriák divergenciáját. Ennek reciprokát nevezzük Ljapunov-időnek, aminek jelentősége ebben a kontextusban az, hogy egy indikátora annak, hogy milyen gyorsan divergálnak a trajektó-

¹egy dimenzió esetén van, hogy egy lineáris rendszerre lehet redukálni azokat, azonban több dimenzióban linearitásukat tekintve irreducibilisek

²példa kaotikus viselkedést generáló paraméterekre: $\sigma = 10$, $\rho = 28$, $\beta = \frac{8}{3}$

riák, így mankóként szolgál majd ahhoz, hogy milyen előrejelzési horizont mellett milyen eredményeket várhatunk.

$$F^t(x_0 + \epsilon) - F^t(x_0) \approx \epsilon \cdot e^{\lambda t}$$
$$t_{\text{Ljapunov}} = \frac{1}{\lambda}$$

1.1.1. Az attraktor

Attraktor alatt a fázistér azon részhalmazát értjük, amely felé egy bizonyos vonzási környezetből kiindulva a dinamikus rendszer aszimptotikusan közelít az idő haladtával. Az attraktort úgy definiáljuk, mint a legkisebb egységet, amely nem bontható fel két vagy több különálló vonzási tartománnyal rendelkező attraktorra. Ez a megszorítás szükséges, mivel egy dinamikus rendszer több attraktort is tartalmazhat, mindegyik saját vonzási tartománnyal.

2. fejezet

Előrejelzés parciális megfigyelésekből

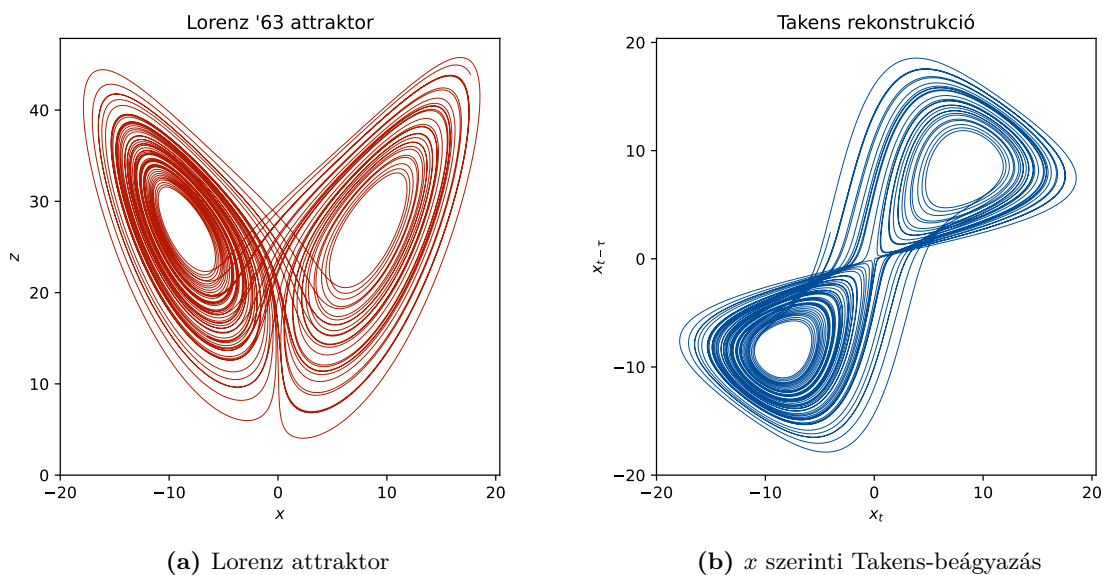
Parciális megfigyelés alatt azt értjük, hogy a rendszerváltozóknak csak egy részét ismerjük. Ebből az is következik, hogy csak ezek közti szabályszerűségeket tudjuk megismerni, így előrejelezni is csak ezeket fogjuk tudni (egyéb adat vagy szabályszerűség ismeretének hiányában). Ebben a dolgozatban egyváltozós megfigyelésekkel fogunk foglalkozni.

2.1. Attraktor-rekonstrukció és Takens tétele

Egy fontos tulajdonsága a vizsgált rendszereknek, hogy minden változó függ a többitől (azaz nem szeparálható a rendszer), így csak akkor tudjuk akármely változó értékét is előrejelezni a következő időpillanatra, ha ismerjük minden változó pontos értékét. Itt jön segítségünkre Takens tétele [19], mely következménye, hogy akárcsak egyetlen megfigyelt változó esetén is rekonstruálni tudjuk az attraktort topológiai izomorfizmus erejéig az alábbi módon:

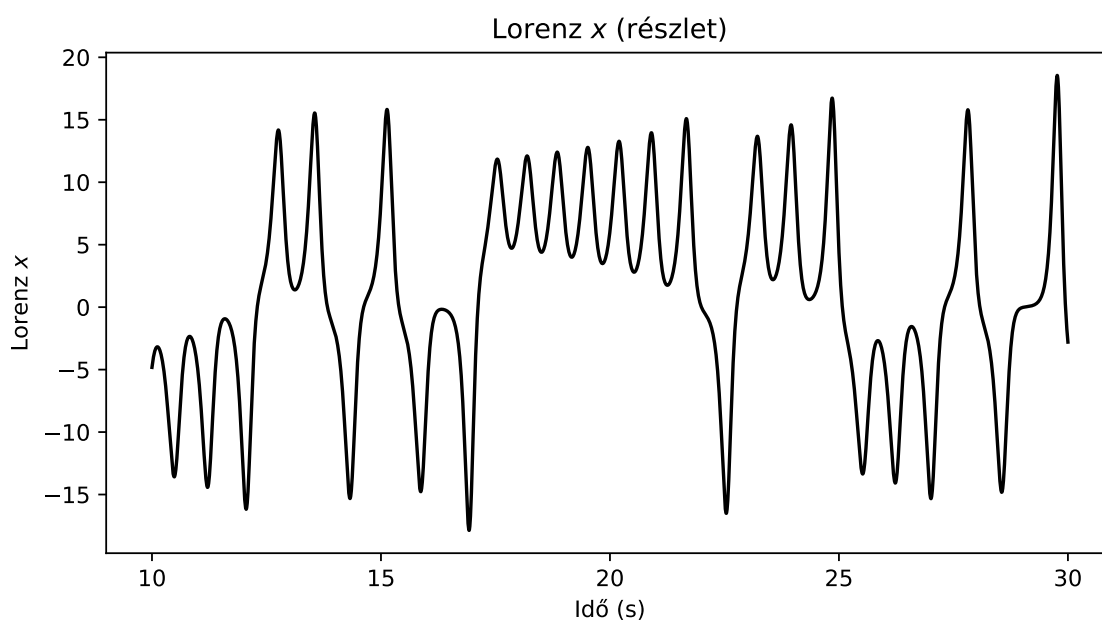
$$\underline{X}_t^{\text{Takens}} = \begin{bmatrix} x_t \\ x_{t-\tau} \\ x_{t-2\tau} \\ \vdots \\ x_{t-(E-1)\tau} \end{bmatrix}$$

ahol x_t az egyik változó a t -edik időpillanatban, $\underline{X}_t$ pedig a rekonstruált attraktor a t -edik időpillanatban. E jelöli a beágyazási dimenziót, τ pedig az eltolás mértékét. A τ meghatározására az átlagos kölcsönös információt használjuk, E -t pedig Cao beágyazási algoritmusával [3] kaphatjuk meg.



2.1. ábra. A Lorenz rendszer x változójának Takens-beágyazása. Minden (x_t, y_t, z_t) ponthoz az attraktor (a) hozzárendelünk egy $(x_t, x_{t-\tau})$ pontot. Bár a keletkezett attraktor (b) az ereditől eltérő formával rendelkezik, jól látszik, hogy a karakterisztikus "pillangó szárnyak" megmaradnak.

Ebben a dolgozatban a Lorenz-rendszer x változójának előrejelzésével fogunk foglalkozni, ezt használjuk összehasonlítási alapként a különböző modellek elbírálásakor. Ezt a választást elsősorban a bizonyított kaotikussága miatt, illetve az irodalomban való gyakorisága miatt hoztuk. Az alábbi ábra egy részletét mutatja a vizsgált idősnak:



2.2. ábra. A Lorenz-rendszer x változójának időszora.

2.2. Kaotikus rendszerek előrejelzése

Kaotikus tulajdonságokat mutató dinamikus rendszerek előrejelzésére több különböző technika is elterjedt. Közös vonásuk, hogy az ismeretlen differenciálegyenleteket próbálják különböző nemlineáris függvényekkel közelíteni, ily módon meghatározva a deriváltakat, amit az előző állapothoz hozzáadva megkapjuk a következő állapotot.

$$\begin{aligned}\nabla_t \underline{X} &= \Phi(\underline{X}_t, \underline{X}_{t-\Delta t}, \dots, \underline{X}_{t-N\Delta t}) \\ \underline{X}_{t+\Delta t} &= \underline{X}_t + \nabla_t \underline{X} \cdot \Delta t\end{aligned}$$

Az egyes megközelítések abban térnek el, hogy a deriváltakat milyen függvénnyel közelítik. Ezen modelleknek van egy fontos részhalmaza: az autoregresszív (AR) modellek. Ezek közös tulajdonsága, hogy a deriváltat az elmúlt elemek valamilyen lineáris kombinációjával határozzák meg:

$$\nabla_t \underline{X} = \phi_0(\underline{X}_t, \underline{X}_{t-\Delta t}, \dots, \underline{X}_{t-N\Delta t}) + \sum_{i=1}^N \phi_i(\underline{X}_t, \underline{X}_{t-\Delta t}, \dots, \underline{X}_{t-N\Delta t}) \underline{X}_{t-i\Delta t}$$

2.2.1. Next Generation Reservoir Computing (NG-RC) [7]

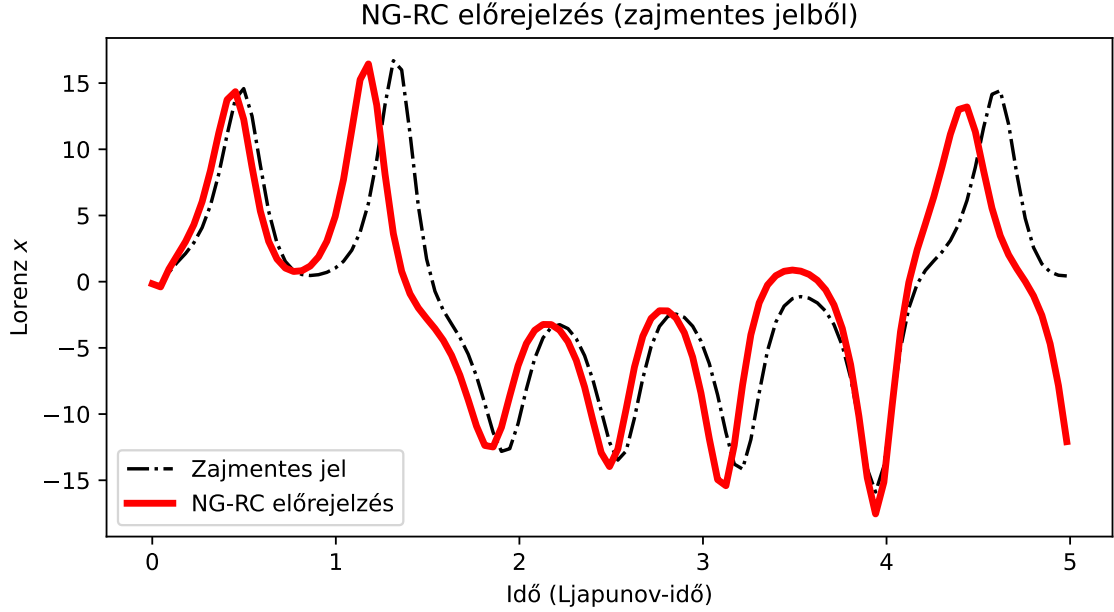
Jelenleg kaotikus rendszerek előrejelzésére az ún. következő generációs tározó számítás a legjobb modell nagyon sok szempontból. A klasszikus Reservoir Computing (RC) [14] modellekkel matematikailag azonos, de számítási hatékonyságát tekintve hatékonyabb NG-RC modellek minden deriváltat polinomokkal közelítenek. Motivációja az, hogy a változók alacsony foksámú polinomjaival már kellően jól tudjuk közelíteni a megtanulni kívánt deriváltakat:

$$\begin{aligned}\underline{Q}_{\text{lineáris}} &= \bigoplus_{i=1}^k \underline{X}_{t-(i-1)s} = \underline{X}_t \oplus \underline{X}_{t-s} \oplus \dots \oplus \underline{X}_{t-(k-1)s} \\ \underline{Q}_{\text{nemlineáris}}^p &= \left[\bigotimes_{i=1}^p \right] \underline{Q}_{\text{lineáris}} \\ \underline{Q}_{\text{nemlineáris}} &= \bigoplus_{p=2}^{P_{\max}} \underline{Q}_{\text{nemlineáris}}^p \\ \underline{Q}_{\text{totál}} &= \left[1 \right] \oplus \underline{Q}_{\text{lineáris}} \oplus \underline{Q}_{\text{nemlineáris}} \\ \underline{Y} &= \underline{X}_{t+\Delta t} - \underline{X}_t \\ \underline{W} &= \underline{Y} \underline{Q}_{\text{totál}}^T \left(\underline{Q}_{\text{totál}} \underline{Q}_{\text{totál}}^T + \alpha I \right)^{-1} \\ \nabla_t \underline{X} &= \underline{W} \cdot \underline{Q}_{\text{totál}}\end{aligned}$$

ahol s , k és $P_{\max}$ hiperparaméterek, $\underline{X}_t$ a rendszer állapotvektorát jelöli egy adott t időpillanatban, $\underline{W}$ a súlymátrix, $\oplus$ a konkatenáció operátort jelöli, és végezetül $\left[\otimes \right]$ jelöl egy olyan operátort, mely egy külső szorzást követően összegyűjti az egyedi polinomokat (a felső háromszögét a keletkezett mátrixnak) és egy vektorba helyezi ezeket. Hatékonysága abban rejlik, hogy ezeket az ún. nemlineáris vonásokat (egynél nagyobb fokú polinomokat)

elég a tanítás során egyszer legenerálni, majd ezekre - a lineáris vonásokkal konkatenálva - lineárisan rátanulni.

Ezek a modellek kiemelkedően jó pontossággal képesek megtanulni az egyes attraktorok dinamikáját, és több Ljapunov-időn keresztül is szinte tökéletes előrejelzéseket képesek szolgáltatni. Kiválóan alkalmazhatóak előrejelzési feladatokra, illetve attraktorok vizsgálatára (pl.: fix pontok keresésére).



2.3. ábra. Az NG-RC modell előrejelzései a zajmentes Lorenz x idősorból.

2.2.2. NG-RC-Chebyshev modell (NG-RC-Ch) [17]

Bár az NG-RC modell nemlineáris jellemzőit alkotó polinomok jól tudják közelíteni a kívánt deriváltakat, egy egészen egyszerű módosítással jelentősen csökkenteni tudjuk a használt polinomok fokát úgy, hogy az előrejelzések pontossága változatlan maradjon (vagy megfordítva: adott fokszám mellett jobb eredményt adjon). Ennek több előnye van, de kiemelkedő az NG-RC algoritmusok rossz térkomplexitása a polinomok fokára nézve. Ha a rendszer állapotvektora n dimenziós, akkor a nemlineáris jellemzők térkomplexitása az alábbi módon alakul:

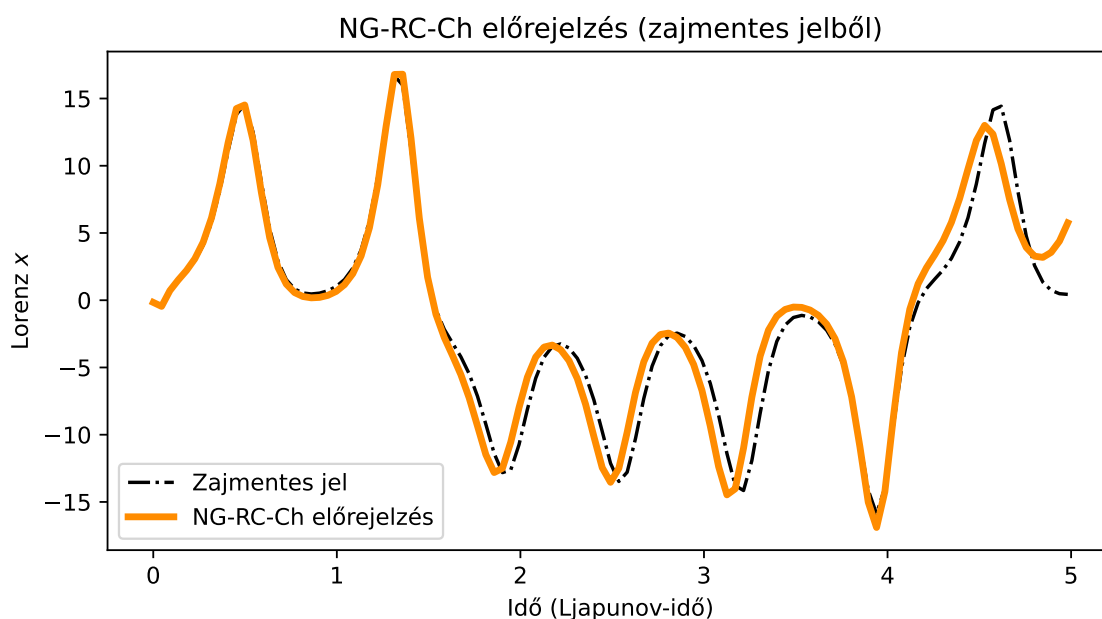
$$\underline{X} \in \mathbb{R}^n \wedge p_{\max} > 1 \implies \underline{Q}_{\text{nemlineáris}} \in \mathbb{R}^m$$

$$m = \binom{n \cdot k \cdot p_{\max}}{n \cdot k} - n \cdot k - 1$$

Emiatt jelentős, hogy csökkenteni tudjuk a fokszám-igényét a modellnek. A gondolat az, hogy minden monomiálist az épített polinomunkban cseréljünk ki Csebisev-polinomokra, amelyeknek az ortogonális tulajdonságát használjuk ki. Tehát a nemlineáris jellemzőkben minden $\underline{Q}_{\text{lineáris},i}^a \cdots \underline{Q}_{\text{lineáris},j}^b$ tagot megfeleltetünk egy $T_a(\underline{Q}_{\text{lineáris},i}) \cdots T_b(\underline{Q}_{\text{lineáris},j})$ tagnak, ahol $T_a(x)$ az a -edik fokú Csebisev-T polinomot jelöli az x helyen.

Bár ezzel a módosítással növekszik az előrejelzési horizont, amin belül magas pontosságot tudunk elérni, mind az NG-RC-nek, mind az NG-RC-Ch modellnek van egy gyengéje:

nehezen küzdenek meg zajos idősorokkal. Minimális zajt tolerálni tudnak a Tikhonov-regularizáció α paraméterének növelésével, azonban a zaj növekedtével (a jel-zaj arány csökkenésével) meglehetősen hamar elromlanak az előrejelzések.



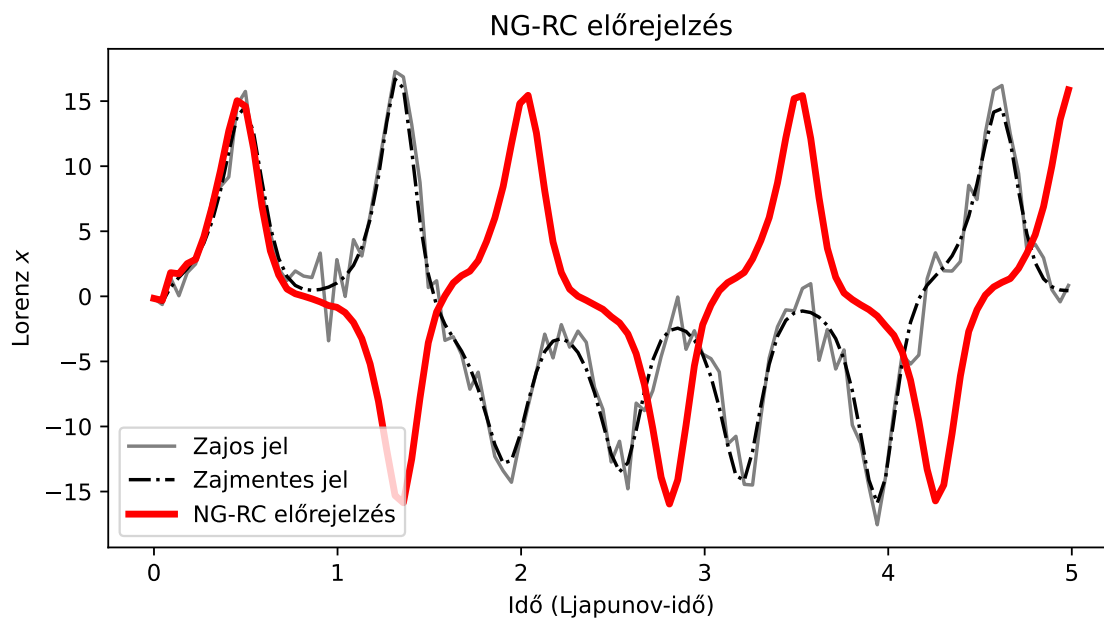
2.4. ábra. Az NG-RC-Ch modell előrejelzései a zajmentes Lorenz x idősorból.

2.3. Előrejelzés zajos megfigyelésekből

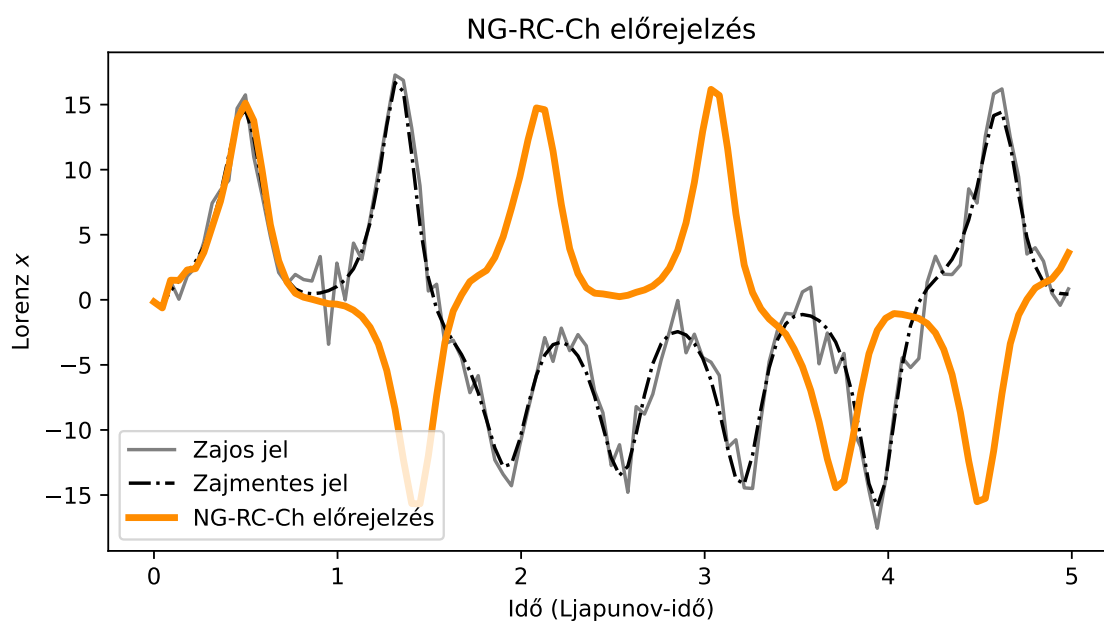
A való életben elkerülhetetlen, hogy a megfigyeléseink zajosak legyenek, mivel a legpontosabb mérőműszer is hibázik, illetve a számítógépes számábrázolás is diszkrét, tehát kvantálást kell alkalmaznunk. Ezek mind megnehezítik az előrejelzést, miután nemcsak a differenciálegyenletrendszer megtanulása válik bonyolultabbá, de a kiindulási állapotot sem ismerjük pontosan.

A dolgozat során ezt a zajt mesterségesen adjuk hozzá a Lorenz idősorhoz (így additív zajt eredményezve). A zaj, amit használunk fehér, a keletkező SNR 14,39 dB. A zaj mértékét úgy választottuk meg, hogy az NG-RC-alapú modellek előrejelzésének pontosságában már látványos romlás legyen megfigyelhető.

A lenti ábrán (2.5) láthatjuk, hogy ez a csökkenés az SNR-ben milyen drasztikus csökkenést jelent az előrejelző képességében az NG-RC-alapú modelleknek. Bár az előrejelzések az attraktor környékén maradnak, gyorsan eltérnek a valós adatsortól. Ebben a dolgozatban ezt a keletkező pontatlanságot szeretnénk csökkenteni egy alternatív modellel.



(a) NG-RC predikciók



(b) NG-RC-Ch predikciók

2.5. ábra. Zajos idősből való előrejelzés az NG-RC és NG-RC-Ch modellekkel. Jól látható, hogy bár az előrejelzés az attraktoron marad (ez abból látszik, hogy alakjában hasonló), az előrejelzés pontatlan (ez abból látszik, hogy 1 Ljapunov-idő után már nagy különbségek vannak az előrejelzett és a valós idősorok közt).

3. fejezet

Empirikus módú dekompozíció (EMD)

Huang 1998-as publikációjának [9] lényege, hogy egy jelet több, úgynevezett alapvető módusú függvényre (IMF - Intrinsic Mode Function) és egy reziduálra bont, melyek összege megegyezik az eredeti jellel. Az IMF-ektől két dolgot várunk el:

- I. Minden két zérushelye közt egyetlen extrémummal kell rendelkeznie.
- II. A zérushelyek és az extrémumok számának különbsége legfeljebb egy lehet.

Ez a felbontás segíti az előrejelzést, mivel különböző frekvenciájú komponensekre bontja azt.

3.1. Az EMD algoritmus

Legyen $x(t)$ a jel, amit fel akarunk bontani. Ekkor az EMD algoritmus lépései a következők:

1. Vegyük a lokális extrémumait $x(t)$ -nek, és ezek közt harmadfokú spline-t illesztünk (a minimumokra és maximumokra külön-külön), így megkapva egy alsó és felső burkolót.
2. Jelöljük az alsó és felső burkolók átlagát $m_1(t)$ -vel.
3. Vonjuk ki az idősből ezt az átlagot: $h_1(t) = x(t) - m_1(t)$.
4. Amennyiben $h_1(t)$ egy IMF (azaz teljesíti a fenti két feltételt), $h_1(t)$ alkotja az eredmény első komponensét - jelölje $c_1(t) = h_1(t)$. Amennyiben $h_1(t)$ nem IMF, úgy ismétljük az eddigi lépéseket (1 - 3), ameddig egy IMF-t nem kapunk.
5. Vegyük az eredeti $x(t)$ idősnak és a kapott $c_1(t)$ komponensnek a reziduálját: $r_1(t) = x(t) - c_1(t)$
6. Helyettesítsük $x(t)$ -t $r_1(t)$ -vel, és ismétljük a korábbi lépéseket (1 - 6), amíg $r(t)$ nem monoton függvény.

Az algoritmus futása után teljesül, hogy az eredeti jelet n komponensre és egy $r_n(t)$ reziduálra bontottuk fel, amelyek szuperpozíciója megegyezik az eredeti jellel.

$$x(t) = r_n(t) + \sum_{i=1}^n c_i(t)$$

Az EMD kiváló eszköz lehet komplex, nemlineáris, akár kaotikus jelek felbontására, mivel ezek az IMF-ek sokszor jobban előrejelezhetőek, mint az eredeti jel.

3.2. Együttes Empirikus Módú Dekompozíció (EEMD) [22]

Az EMD nagy sikert aratott, és széles körben alkalmazták, azonban több hiányossággal is rendelkezik, köztük azzal, hogy az egyes komponensek spektrumai keverednek, azaz nem szeparálódnak el a frekvencia-tartományok. Ez számunkra előnytelen, mivel így nehezebb előrejelezni. Ezt a gyengeséget hivatott az EEMD (Ensemble Empirical Mode Decomposition) orvosolni, amely minden lokális extrémum keresés előtt fehér zajt ad a jelhez, majd így bontja fel és kapja meg az IMF-eket és a reziduált. Az egész algoritmust többször (általában több százszor, ezerszer) megismétli (mindig más zajjal), majd az eredményül kapott IMF-eket átlagolja, így megkapva a végleges felbontást. Erre azért van szükség, mert az eredeti EMD algoritmus érzékeny a mintavételi frekvenciára, illetve sokszor nem különülnek jól el a különböző frekvenciájú tagok a különböző komponenseken (azaz keverednek a komponensek spektrumai).

Módosítások az EMD algoritmushoz képest:

- Beszúrjuk az 1-es pont elé az alábbi lépést: $x(t) := x(t) + n(t)$, ahol $n(t)$ a generált zaj.
- Az egész algoritmust elvégezzük M -szer, így $M \cdot N$ számú komponenst kapva, majd komponensenként átlagolunk:

$$c_i(t) = \frac{1}{M} \sum_{j=1}^M c_{i,j}(t),$$

ahol $c_{i,j}(t)$ a j -edik futtatás által eredményezett i -edik komponenst jelöli.

Az EEMD által eredményül adott IMF komponensek előrejelezhetősége jobb, mivel az eredeti zaj jobban elkülönül a legmagasabb frekvenciájú komponensekre, így a többin szignifikánsan kisebb lesz a hatása.

3.3. Teljes EEMD (CEEMD) [23]

Az EEMD jól megoldja a spektrumok összekeveredésének problémáját, azonban teremt egy újat: az eredményül kapott komponensekben az átlagolás után is marad némi zaj. Ez jellemzően nem túl nagy, ha M (az algoritmus futása során generált teljes dekompozíciók száma) elég nagy, azonban szeretnénk M -et minimalizálni annak érdekében, hogy minél kevesebb iterációban, minél gyorsabban le tudjuk futtatni az algoritmust. Itt jön a képbe a CEEMD (Complete Ensemble Empirical Mode Decomposition) [23], amely csupán annyit módosít az EEMD algoritmuson, hogy minden általunk hozzáadott zajhoz veszi annak ellentettjét, és azzal is elvégzi a felbontást, hogy ezek a hozzáadott zajok az átlagolás során páronként kiejtsek egymást, így tisztább komponenseket kapunk a végén.

3.4. Adaptív zajos CEEMD (CEEMDAN) [20]

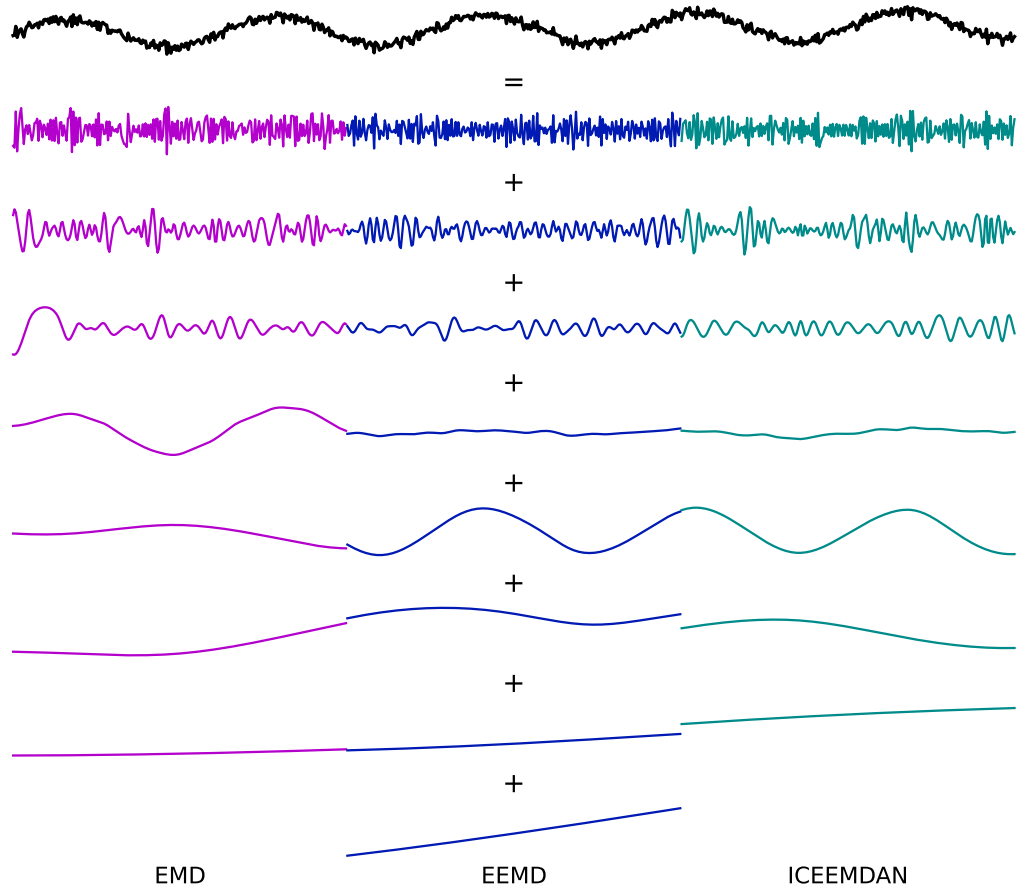
Egy újabb fejlesztés a dekompozíciós technikán az ún. adaptív zajos CEEMD (CEEMD Adaptive Noise), A gondolat az, hogy az iterációk számát és a spektrum szeparációját segíti, ha nem külön-külön, függetlenül futtatunk sok EEMD-t, hanem először sok, páronként ellentétes zajt adunk a jelhez, majd ezeknek mind megkeressük az első IMF komponensét (csakúgy, mint a CEEMD esetében). Most azonban nem csak a lelegején adjuk hozzá a zajt (az eredeti jelhez), hanem minden egyes lépésnél vesszük a reziduálját, és szuperponáljuk egy új fehér zajjal. Ennek eredménye, hogy még jobban elkülönülnek az egyes komponensek spektrumai.

3.5. Továbbfejlesztett CEEMDAN (ICEEMDAN) [6]

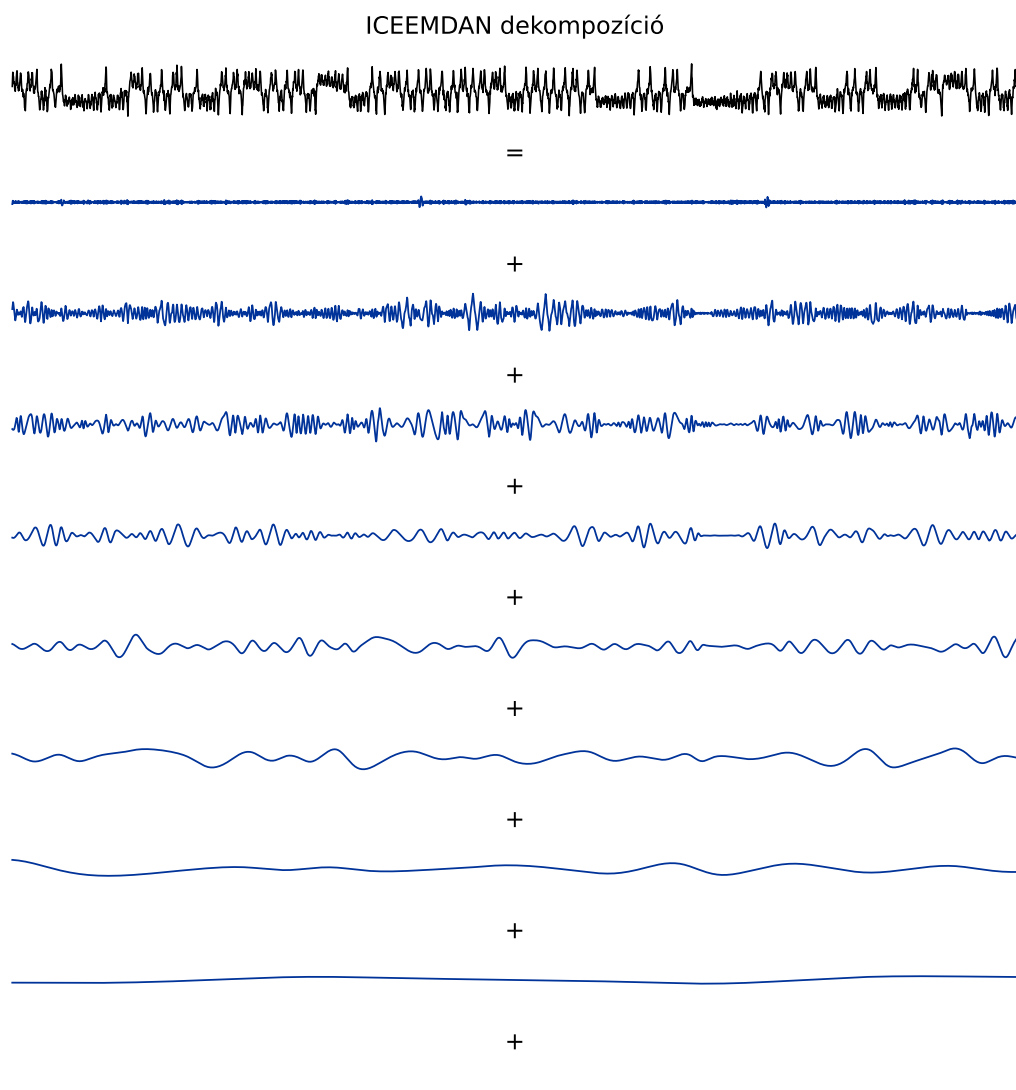
Bár a CEEMDAN algoritmus lényegesen kevesebb reziduált állít elő, mint a sima CEEMD, és széles körben is alkalmazzák, akad pár aspektus, amiben lehet rajta javítani: marad a komponensekben egy kis zaj (az eredeti jelből), illetve a fontosabb komponensek (amik analízishez, vagy előrejelzéshez lennének használva) később jelennek csak meg, és az első pár komponens sokszor zajos és nem releváns a predikció szempontjából. Ezeket a problémákat hivatott a továbbfejlesztett CEEMDAN (Improved CEEMDAN) [6] megoldani úgy, hogy az egyes reziduálokhoz hozzáadott zajt folyamatosan csökkenti, így jазtalanítva az eredő komponenseket, és szeparálva még jobban a komponensek spektrumait.

Ebben a dolgozatban az ICEEMDAN algoritmust fogjuk alkalmazni annak érdekében, hogy minél jobban szeparált komponenseket kapjunk, és minél könnyebb dolga legyen az előrejelző modellnek. Alapvetően elmondható, hogy minél szűkebb spektrumú komponenst kell előrejelezni, annál könnyebb dolgunk van (szélsőséges esetben egy egyszerű szinuszoidot kapunk).

Példa az EMD, EEMD és ICEEMDAN algoritmusok összehasonlítására



3.1. ábra. Az EMD, EEMD és ICEEMDAN algoritmusok összehasonlítása egy egyszerű idősoron. Az idősort a következőképpen kaptuk meg: $x(t) = \sin 10t + 0,2t + \epsilon(t)$, ahol $\epsilon(t) \sim \mathcal{N}(0; 0, 2^2)$. Minden technikával 1-1 másodpercnnyi felbontást látunk. Az egyes komponensek jobb láthatósága érdekében a skálázás változik a komponensek között! Érdemes megfigyelni az EMD szinuszos komponensének amplitúdóváltozását (ami részben a véghatás következménye), illetve zajosságát. Ezzel szemben a másik kettőn véghatást nem tapasztalunk, a szinusz egyenletes amplitúdóval rendelkezik, azonban az EEMD eggyel több komponenst eredményezett (mivel a lineáris részt kettébontotta). Az ábrán szintén jól látszik az ún. véghatás (az EMD felbontás elején).



3.2. ábra. Az ICEEMDAN dekompozíció eredménye a teljes zajos Lorenz- x idősoron. Jól látszik, hogy egyre alacsonyabb frekvenciájú komponensekre bont, illetve hogy elég jól szeparálódnak a spektrumok.

4. fejezet

Radiális Bázisfüggvény Hálózatok

4.1. Radiális Bázisfüggvény Hálózat (RBFN)

Ahhoz, hogy az egyes komponenseket előrejelezzük, mindegyiknek fel kell térképeznünk az attraktorát. Ilyen nemlineáris attraktorok előrejelzésére széles körben elterjedt [4] [8] a radiális bázisfüggvény hálózaton alapuló előrejelzés.

Az eredetileg numerikus interpolációs technikának szánt radiális bázisfüggvény hálózaton [1] alapuló RBF-N (Radial Basis Function Network) modell [16] az egyes deriváltakat RBF-interpolációval közelíti, azaz minden derivált egy-egy lineáris kombinációja lesz az aktuális rendszerállapot $\underline{X}_t$ és több középpont $\{\underline{C}_1, \underline{C}_2, \dots, \underline{C}_n\}$ Gauss-kernel szerint vett távolságának.

$$\nabla_t \underline{X} = \underline{w}_0 + \sum_{k=1}^n w_k \exp \left(-\frac{\|\underline{X}_t - \underline{C}_k\|_2^2}{2\sigma_k^2} \right)$$

4.1.1. K-Means klaszterező algoritmus [13]

A középpontokat sokféleképpen kiválaszthatjuk, azonban sokszor a legcélszerűbb klaszterezni annak érdekében, hogy minél kevesebb középponttal (klaszterrel) az eredeti teret jól lefedve csoportosíthassuk. A választásunk a K-Means [13] módszerre esett, miután ez egy olyan klaszterezési algoritmus, amely az egyes klasztereken belüli pontok és a hozzájuk tartozó (legközelebbi) klaszter középpontjának euklideszi távolságának négyzetét minimalizálja. Legyen $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ a klaszterezni kívánt adathalmaz, ezeket klaszterezzezzük $S_1, S_2, \dots, S_k$ halmazokba. Ekkor az alábbi összefüggések szerint optimalizál a K-Means:

$$\begin{aligned} \mathbf{S}_{\text{optimális}} &= \underset{\mathbf{S}}{\operatorname{argmin}} \sum_{i=1}^k \sum_{\underline{x} \in S_i} \|\underline{x} - \underline{\mu}_i\|^2 \\ \mathbf{S} &= \{S_1, S_2, \dots, S_k\} \\ \forall i \neq j \in [1, 2, \dots, K] : S_i \cap S_j &= \emptyset \\ \mathbf{S} &= \bigcup_{i=1}^K S_i \\ \underline{\mu}_i &= \frac{1}{|S_i|} \sum_{\underline{x} \in S_i} \underline{x} \end{aligned}$$

4.2. Autóregresszív RBFN (RBF-AR) [12]

Az egyszerű RBFN hálózatoknak egy továbbgondolt változata az ún. RBF-AR vagy AR-RBF (Autoregressive RBFN), amely nem csupán a középpontoktól való Gauss-távolság lineáris kombinációját tekinti a derivátnak, hanem minden középponthez rendelünk egy AR-modellt, és ezeknek az előrejelzéseit súlyozzuk az adott középponttól való távolsággal.

$$\begin{aligned}\nabla_t \underline{X} &= \underline{\phi}_0 + \sum_{i=1}^m \underline{\phi}_i \odot \underline{X}_{t-i+1} \\ \underline{\phi}_0 &= \underline{w}_{0,0} + \sum_{k=1}^n \underline{w}_{0,k} \exp\left(-\frac{\|\underline{X}_t - \underline{C}_k\|_2^2}{2\sigma^2}\right) \\ \underline{\phi}_i &= \underline{w}_{i,0} + \sum_{k=1}^n \underline{w}_{i,k} \exp\left(-\frac{\|\underline{X}_t - \underline{C}_k\|_2^2}{2\sigma_k^2}\right)\end{aligned}$$

Itt az autóregresszív előrejelzések visszatekintési ablakának méretét m jelöli, $*$ pedig a vektorok komponensenként való szorzását reprezentálja (azaz $\begin{bmatrix} 2 \\ 3 \end{bmatrix} \odot \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 8 \\ 15 \end{bmatrix}$).

Ha intuitívan szeretnénk megmagyarázni, hogy mit csinál a modellünk, akkor azt mondhatjuk, hogy minden középponthez megtanul egy AR-folyamatot, amik előrejelzésének lineáris kombinációját veszi. Ezen lineáris kombináció együtthatóit az egyes középpontoktól való távolság határozza meg (minél közelebb vagyunk a ponthez, annál dominánsabb lesz a hatása).

4.3. Fuzzy RBF-AR (FRBF-AR) [10]

Fuzzy logikában az igazságértékek a $[0; 1]$ tartományban akármilyen valós értéket felvehetnek (0 a hamis, 1 az igaz állítást jelöli). A konjunkció és diszjunkció műveleteknek egy lehetséges fuzzy megfelelője a szorzás és összeadás műveletek. Ennek ismeretében alakítsuk át az RBF-AR modellt a következőképpen:

1. Legyen ψ_k annak az igazságértéke, hogy az attraktorban közel vagyunk a C_k ponthez.
2. Vegyük ezeknek az állításoknak a konjunkcióit (jelölje ν_j) oly módon, hogy minden eredő állítás r korábbi állítás konjunkciójából álljon. Ezt kifejezhetjük úgy, hogy $\nu_j = \prod_{k \in K_j} \psi_k$, ahol K_j jelöli az j -edik konjunkciót képező állítások indexeinek halmazát ($|K_j| = r$). Nevezzük ezeket a konjunkciókat jellemzőknek.
3. Minden jellemzőhöz rendeljünk egy AR-modellt (együtthatókat), és súlyozzuk ezek előrejelzéseit aszerint, hogy melyik jellemző (konjunkció) teljesül. Ez a súlyozott előrejelzés lesz az előrejelzésünk.

Ezekkel a módosításokkal az alábbi modellt kapjuk eredményül:

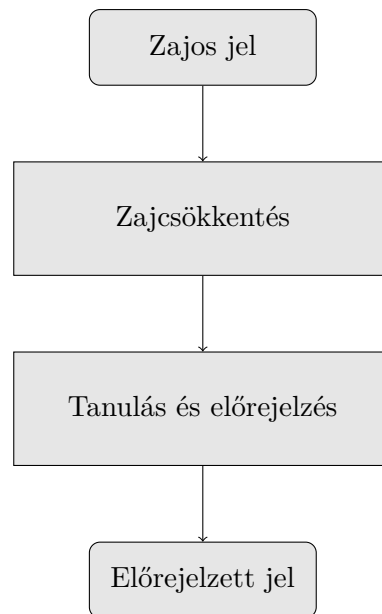
$$\begin{aligned}
\nabla_t \underline{X} &= \underline{\phi}_0 + \sum_{i=1}^m \underline{\phi}_i * \underline{X}_{t-i+1} \\
\underline{\phi}_0 &= \underline{w}_{0,0} + \sum_{j=1}^n \underline{w}_{0,j} \nu_j \\
\underline{\phi}_i &= \underline{w}_{i,0} + \sum_{j=1}^n \underline{w}_{i,j} \nu_j \\
\nu_j &= \prod_{k \in K_j} \phi_k \\
\phi_k &= \exp \left(-\frac{\|\underline{X}_t - \underline{C}_k\|_2^2}{2\sigma_k^2} \right)
\end{aligned}$$

Fontos megjegyezni, hogy ennél a modellnél klaszterezéssel kapjuk meg a kezdeti kö-zéppontokat, azonban ezek (a σ szórásokkal együtt) tanulható paraméterek. Ez a korábbi modelleknél is lehet így, azonban az FRBF-AR modellek ezek tanulása nélkül jelentősen rosszabbul extrapolál (míg a többinél nem ilyen látványos ez a csökkenés az előrejelző képességben).

5. fejezet

Az ICEEMDAN-FRBF-AR modell

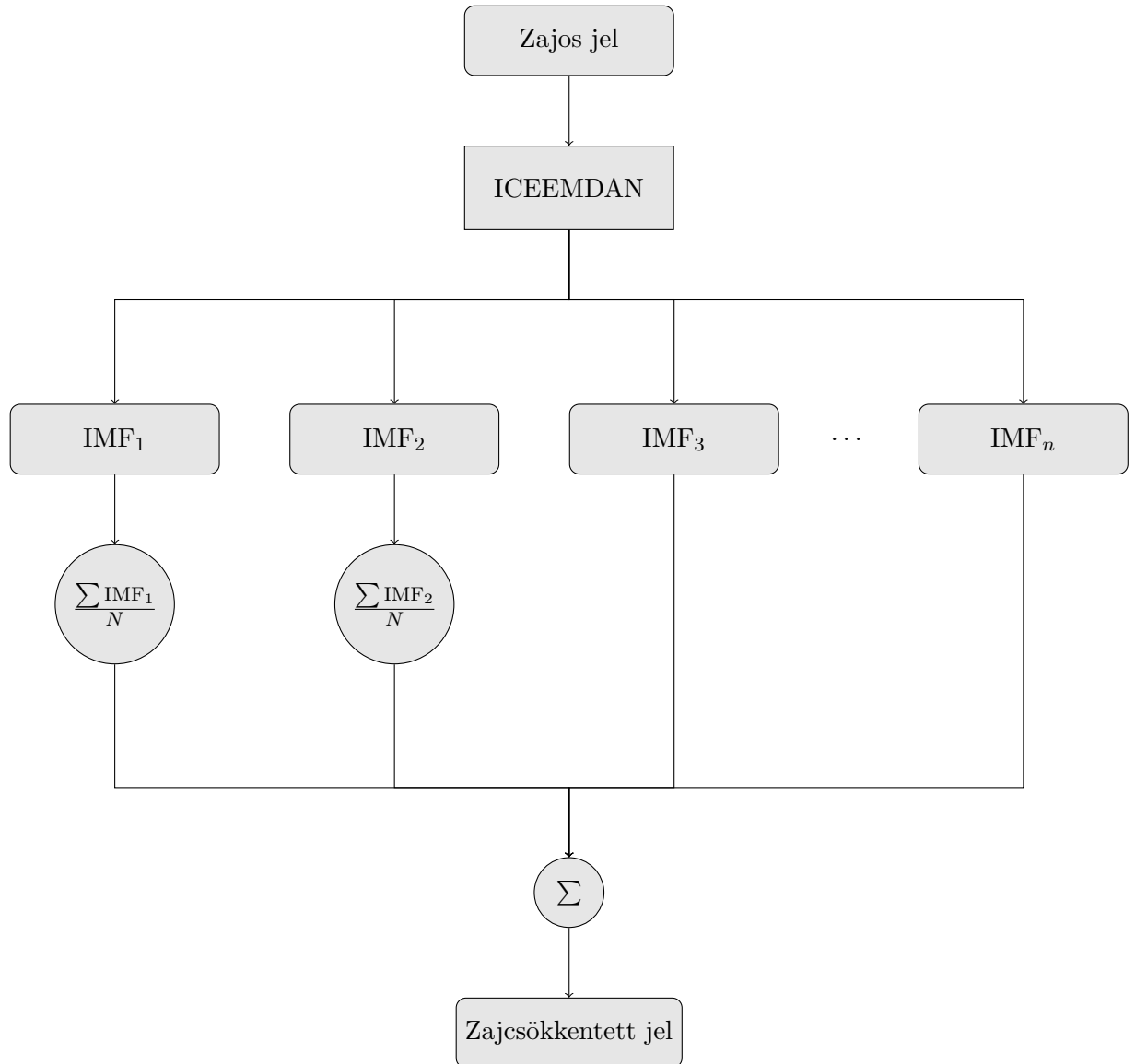
Ezen dolgozatban javasolt modell a zajos, kaotikus parciális megfigyelésekből való előrejelzésre az ICEEMDAN-FRBF-AR modell, amely az alábbi struktúrával rendelkezik:



5.1. ábra. ICEEMDAN-FRBF-AR modell magas szintű ábrája

5.1. Zajcsökkentés

Miután az ICEEMDAN által eredményezett IMF komponensek spektrumai jól elkülönülnek, és a megfigyelési zajunk frekvenciája igen magas, a legmagasabb frekvenciájú komponenseken "kigyűlik" a zaj, azaz ezeket elhagyva tisztább jelet kaphatunk. Az elhagyható komponensek száma egy hiperparaméter, azonban tapasztalataink (5.1) azt mutatják, hogy 1-2-nél több komponens elhagyásával sok információt veszünk a jelről. Itt is segítségünkre van, hogy mesterségesen adtuk a tiszta jelhez a zajt, miután így pontosan ki tudjuk számolni az SNR-t (Signal-to-Noise ratio) - valós adatsorok esetén ezt csak közelíteni lehet. Erre a közelítésre jó megoldás lehet az RC (vagy akár NG-RC) modellekkel való interpoláció [5].



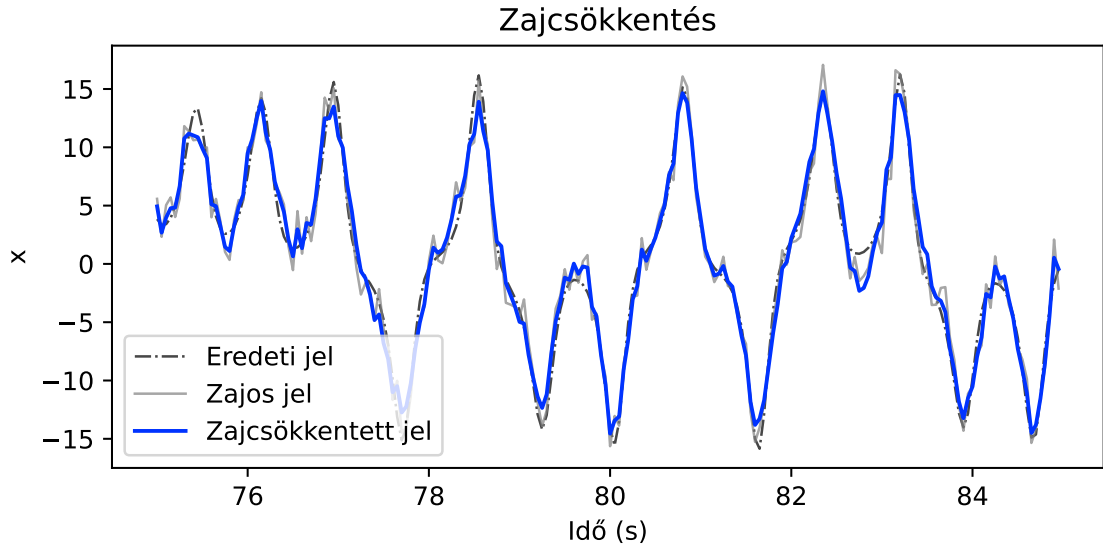
5.2. ábra. Zajcsökkentés 2 IMF komponens elhagyása esetén.

Elhagyott IMF-ek száma	SNR (dB)	R ²
0	14,39	0,964
1	14,398	0,964
2	14,892	0,968
3	7,588	0,829

5.1. táblázat. Jel-zaj arány a zajcsökkentés előtt és után. Az SNR mellett feltüntettük az R²-et, ami a determinisztikussági együtthatót jelöli.

5.2. Tanulás és előrejelzés

A kulcsgondolat a modellben az, hogy a zajos, nemlineáris jelet bontsuk fel az ICEEMDAN segítségével több, egyszerűbb jelre, amelyeket külön-külön egyszerűbb előrejelezni. Ezeket



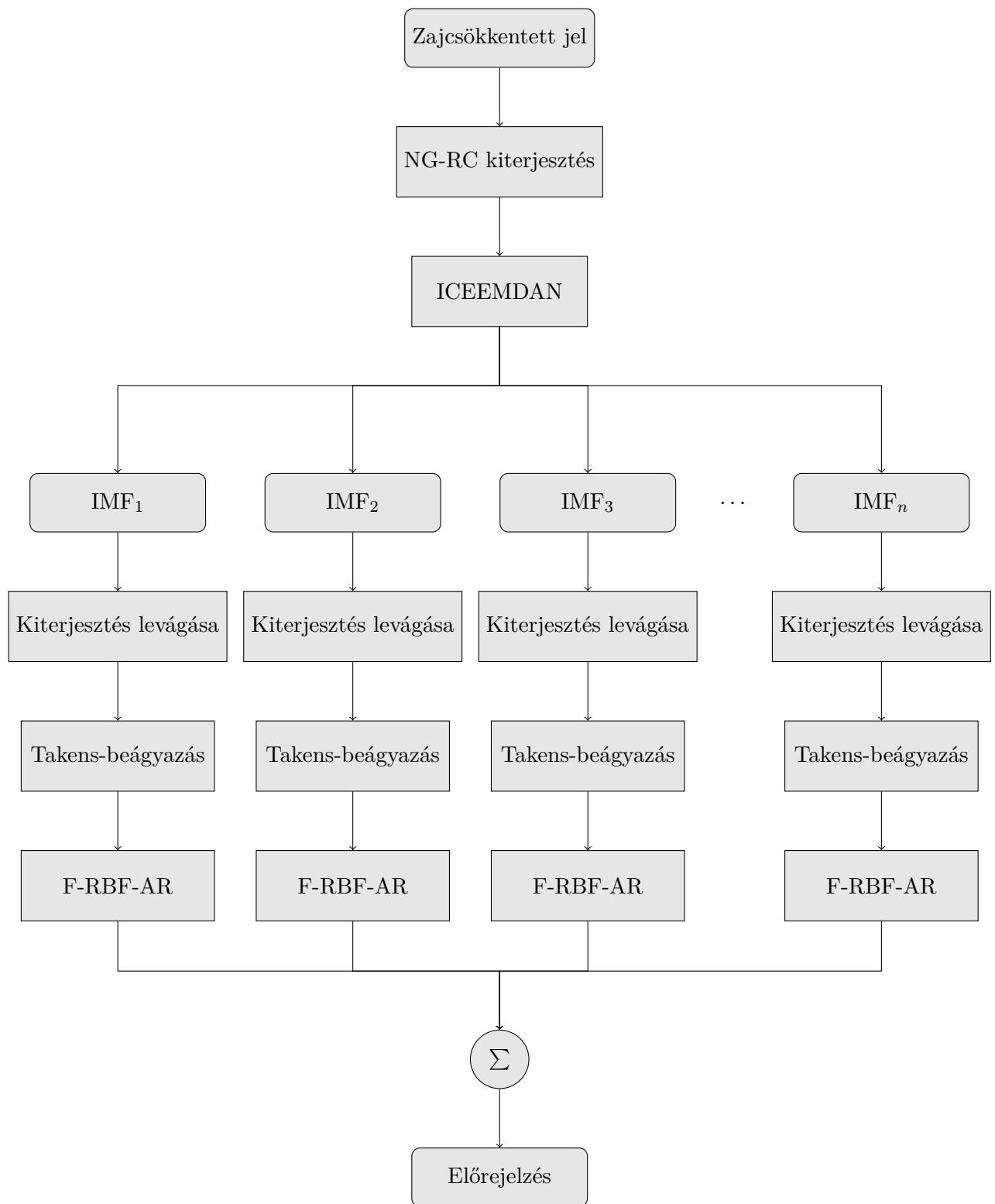
5.3. ábra. *A zajcsökkentett jel részlete.* Jól látszik, hogy bár a magas frekvenciájú zajokat jól levágja, és az SNR-t javítja, a zajcsökkentés a magasfrekvenciájú IMF-ek elhagyása az eredeti jelet is torzítja, így ezekkel a tanuló algoritmusnak kell megküzdenie.

egy RBF-N modellel közelítjük [16], majd az előrejelzett IMF-eket összeadjuk, megkapva a végleges előrejelzést.

A felbontott IMF-eket Takens-beágyazással beágyazunk egy magasabb dimenziójú térbe, így megkapva minden IMF-hez a saját fázis-terét és benne attraktorát.

Az IMF-ekre való dekompozíciónak az is nagy előnye, hogy amennyiben vannak olyan IMF-ek, amelyeket nehezen tudunk megtanulni, és esetleg az előrejelzés csak egy oszcilláló szinuszoid lesz, vagy akár beáll egy konstans értékre, akkor is a többi IMF által nyújtott előrejelzés elég jó közelítése lehet az eredetinek. Emiatt az előrejelzések stabilitása lényegesen jobb lesz, mint a konkurens modelleké (NG-RC, NG-RC-Ch). Ennek háttérében az áll, hogy míg teljes részletességében nehéz előrejelezni a jelet, addig csak egyes részeit a spektrumnak kellő pontossággal közelíteni tudjuk. Itt többek között szerepet a visszatérési ablak mérete, illetve a modell paramétereinek száma - minél magasabb frekvenciájú komponenst akarunk előrejelezni, annál komplexebb modellre van szükségünk, és magas E és kis τ paraméterekkel való beágyazásra. A gyakorlatban azt tapasztaljuk, hogy a magas frekvenciájú IMF-ek előrejelzése hamarabb eltér a valóstól, mint az alacsonyabbaké, ami intuitív is, mivel minden "részlete" a jelnek, illetve a megmaradt zaj egy jó része a magas frekvenciákban lesz fellelhető.

Összességében elmondhatjuk, hogy még ha a magas frekvenciájú IMF-ek által szolgáltatott részletességet el is veszítjük az előrejelzési ablak növelése során, a maradék IMF egy kellően jó képet ad ahhoz, hogy összességében még elég pontos előrejelzést kapjunk a modelltől.



5.4. ábra. Az előrejelző modell magas szintű ábrája.

5.2.1. Az F-RBF-AR modell Takens-beágyazás mellett

Amennyiben F-RBF-AR modellel szeretnénk előrejelezni egy egydimenziós változót, az alábbi módon kell módosítani a korábbi felírást:

$$\begin{aligned}
\dot{x}(t) &= \frac{x_{t+1} - x_t}{\Delta t} = \phi_0 + \sum_{i=1}^m \phi_i \cdot x_{t-i+1} \\
\phi_0 &= w_{0,0} + \sum_{j=1}^n w_{0,j} \nu_j \\
\phi_i &= w_{i,0} + \sum_{j=1}^n w_{i,j} \nu_j \\
\nu_j &= \prod_{k \in K_j} \phi_k \\
\phi_k &= \exp \left(- \frac{\| \underline{X}_t^{\text{Takens}} - \underline{C}_k \|_2^2}{2\sigma_k^2} \right) \\
\underline{X}_t^{\text{Takens}} &= \begin{bmatrix} x_t \\ x_{t-\tau} \\ x_{t-2\tau} \\ \vdots \\ x_{t-(E-1)\tau} \end{bmatrix}
\end{aligned}$$

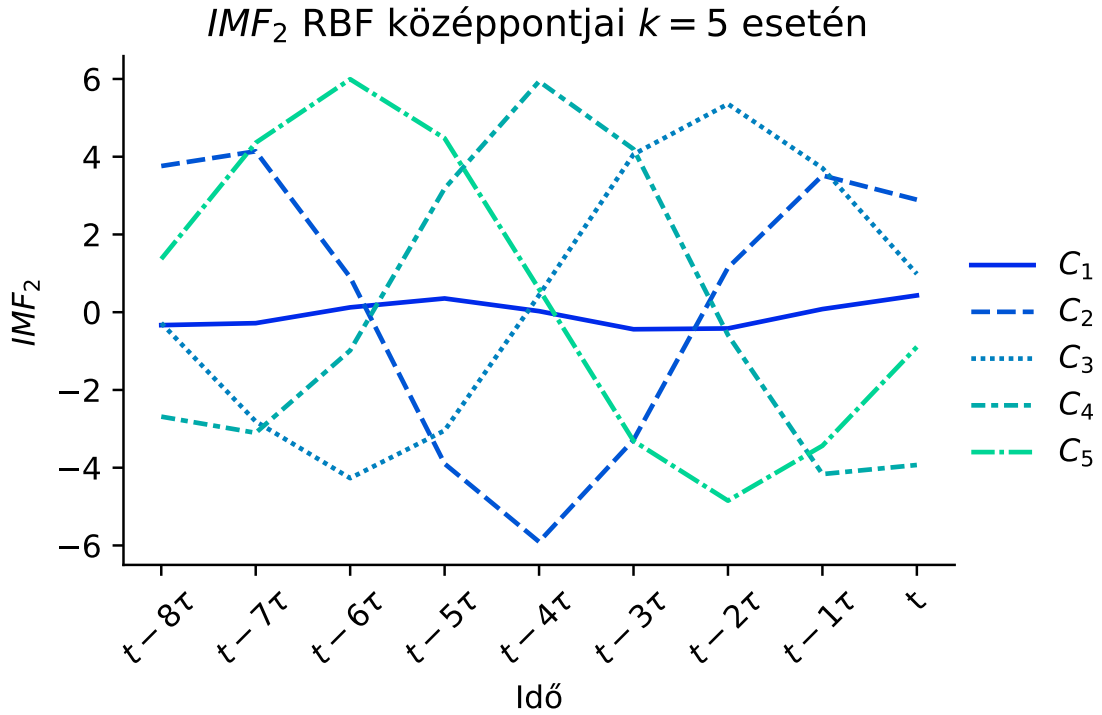
ahol K_j azon indexek halmaza, amelyek a j -edik jellemzőt alkotják.

A legfontosabb különbség, hogy nem a teljes gradienst közelítjük, hanem csak egy egydimenziós deriváltat, mivel a Takens-beágyazás következtében az első komponens értéke, és az idősor eltoltsjai meghatározzák a többi komponenst.

5.2.2. Intuíció többféleképpen

Amennyiben intuíciót szeretnénk nyerni, fogalmazhatunk úgy, hogy az F-RBF-AR elhelyez a tanuló attraktorba pontokat (ezek lesznek az RBF középpontok), amiknek konjukcióihoz (azaz azokhoz a pontokhoz, amik az adott jellemzőt alkotó középpontokhoz közel vannak) asszociál egy előrejelző (AR) folyamatot. Előrejelzéskor az a tanult AR folyamat fog dominálni, ami a legrelevánsabb a rendszer aktuális állapotában - azaz amelyik metszetbe (konjukcióba) a leginkább besorolható a rendszerállapot t időpillanatban.

Amennyiben a rendszer állapotvektorát Takens-beágyazással kapjuk meg (mint ahogyan az ICEEMDAN-F-RBF-AR modellnél is), felmerül egy alternatív magyarázata annak, hogy hogyan is tanul a modellünk: miután minden egyes pont az idősor utolsó pár (E) időpontjából áll (τ eltolással), gondolhatunk ezekre a pontokra úgy, mint az idősor végének lenyomatára, amiket a klaszterezés során homogén halmazokba sorolunk, amiknek a folytatása nagyon hasonló, és jól közelíthető egy AR folyamattal. Előrejelzésnél megnézzük, hogy melyik ilyen homogén csoporthoz mennyire hasonlít (RBF távolság), majd ezzel súlyozzuk az egyes homogén csoportokhoz tartozó folytatást. Az 5.5 ábrán látható példában, ahol $k = 5$ klaszterbe soroltunk egy IMF-nek a Takens-beágyazott idősorát. Jellemzően jóval több klaszterbe sorolnánk a pontokat, azonban gyorsan követhetetlen lenne az ábra, így az átláthatóság érdekében limitáltuk az RBF középpontok számát. Itt nem tüntettük fel a jellemzőket, csupán a középpontokat, azonban a jellemzők az ezektől vett távolságok szorzatai.



5.5. ábra. A beágyazott IMF_2 idősor klasztereinek értelmezése 5 klaszter esetén. Jól látszik, hogy mennyire eltérnek egymástól, így minimalizálva a multikolinearitást a tanulás során. Bár jellemzően 30-300 RBF-el szoktunk közelíteni, az átláthatóság végett 5 középpontot szemléltetünk.

5.2.3. NG-RC kiterjesztés

Miután korábban láttuk, hogy az NG-RC-alapú modellek rövid távon jó pontossággal tudnak előrejelezni, még az ICEEMDAN felbontás előtt egy NG-RC előrejelzéssel "ki tudjuk terjesztetni" a jelet előre és hátra (1-1 Ljapunov-időre). Ez azért fontos, mert így minimalizálni tudjuk az ún. véghatás-effektusát az ICEEMDAN-nak.

Véghatás alatt azt értjük, amikor az IMF komponensek eleje és vége más tulajdonságokkal bír, mint a köztes rész (pl.: hirtelen csökken a frekvencia a végekhez közel érve). Ez a hatás jól látható a 3.1 ábrán, ahol az EMD felbontás elejét az eredeti idősor elején kezdtük. Látható, hogy vannak komponensek, amik más viselkedést mutatnak az első pár tized másodpercben, mint később. Ez az előrejelzésre igen erős hatással van, mivel mi az egyes komponensek végét szeretnénk folytatni előrejelzésre, így kritikus, hogy ezen komponensek ne változzanak meg a végeiken.

Miután az ICEEMDAN dekompozíciót elvégeztük, nincs már többé szükségünk a kiterjesztésekre, így egyszerűen levágjuk őket (így az eredeti idősor komponensei maradnak csak, viszont véghatás nélkül).

5.2.4. Hiperparaméterek és megválasztásuk

A modell több hiperparaméterrel is rendelkezik, amelyeket heurisztikus módon, illetve különböző hiperparaméter-optimalizáló algoritmussal is optimalizálhatunk (pl.: bayesi modellek, rács keresés, stb.), ezeket taglaljuk röviden a következőkben.

5.2.4.1. Zajcsökkentéskor elhagyott IMF-ek száma

A zajcsökkentés során az ICEEMDAN által adott IMF-ek közül a legnagyobb frekvenciájúakat elhagyjuk, így redukálva a zajt, azonban felmerül a kérdés, hogy mennyit is hagyjunk el? Természetesen a jel-zaj arányt akarjuk maximalizálni, de honnan tudjuk, hogy az mennyi? A jel-zaj arány becslésére több heurisztika is létezik [2] [5], azonban a vizsgált esetben (miután a zajt mi adtuk a vizsgált idősorhoz) pontosan meg is tudjuk válaszolni a kérdést. A vizsgált esetben az eredmények 5.1 igazolják, hogy egy pár IMF elhagyása (esetünkben pontosan 2) esetén maximalizáljuk a jel-zaj arányt. Általánosságban elmondható, hogy amennyiben az eredeti jel magasfrekvenciás IMF-jeinek kicsi a jelteljesítménye, ott ezen IMF-ek elhagyása kis torzítást okoz (ugyanakkor nagy zajcsökkenést).

5.2.4.2. RBF középpontok száma

A középpontok számának megválasztása szorosan összefügg a középpontokhoz tartozó sugár nagyságával. Összességében elmondható, hogy minél komplexebb a jel, annál részletesebben akarjuk feltérképezni az attraktort, azaz annál több középpontot szeretnénk (K-Means k paraméter). Egyszerűbb jelek esetén elegendő pár tucat középpont, azonban ez több százig is felmehet egy komplex jel esetén. Fontos, hogy mindig a legkevesebb középpontot szeretnénk, amivel még jól tudunk közelíteni, hogy a modellünk minél jobban tudjon extrapolálni, minimalizálva a túlilleszkedés kockázatát.

5.2.4.3. Visszatekintési ablak az AR folyamatoknál (m)

Az m megválasztásához nehéz jó mankót adni. A mi modellünkben tanítás közben megfeleltettük az attraktor beágyazási dimenziójának ($m = E$), azonban ez csak egy heurisztika, és további hiperparaméter-optimalizáció mindenképpen szükséges optimális értékek eléréséhez.

6. fejezet

Eredmények

6.1. A vizsgált adathalmaz

A modellt a Lorenz rendszer x változójának előrejelzésén teszteljük, amihez fehér zajt adtunk (a kezdeti SNR 14,39dB). A mintavételi frekvencia 20Hz (azaz $\Delta t = 0,05s$), az egész adatsor 200s hosszú, aminek első 194,5 másodpercén tanítjuk, az utolsó 5,5 másodpercen (5 Ljapunov-időn) teszteljük (1 Ljapunov-idő = 1,104 másodperc). A Lorenz rendszer kiintegrálása során az első 20 másodpercet levágtuk (így 200s marad), hogy adjunk időt a rendszernek arra, hogy az attraktorra kerüljön.

6.2. Előrejelzések pontosságainak számszerűsítése

6.2.1. Determinációs együttható

Idősorok előrejelzéseinél egy bevett mutató a determinációs együttható, amire sokszor csak R^2 -ként hivatkoznak. Értéke $] - \infty; 1]$ tartományba esik, ahol 1 a tökéletes egyezés (azaz minél nagyobb az érték, annál jobb).

$$R^2 = 1 - \frac{\sum_{i=1}^n y_i - \hat{y}_i}{\sum_{i=1}^n y_i - \bar{y}}$$
$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

A megfigyelt (valós) adatokat y , míg az előrejelzetteket $\hat{y}$ jelöli.

6.2.2. Átlagos Arkusztangens Abszolút Százalékos Hiba (MAAPE) [11]

Egy újabb széles körben alkalmazott metrika (noha talán kissé kevésbé ismert) az ún. MAAPE (Mean Arctangent Absolute Percentage Error), ami az elterjedtebb MAPE (átlagos abszolút százalékos hiba) egy továbbgondolt változata. Alkalmazásának előnye, hogy a 0-közi értékekre nem ad kiugróan magas hibákat [11]. Értéke a $[0; \infty[$ tartományba esik, minél kisebb, annál jobb az előrejelzés.

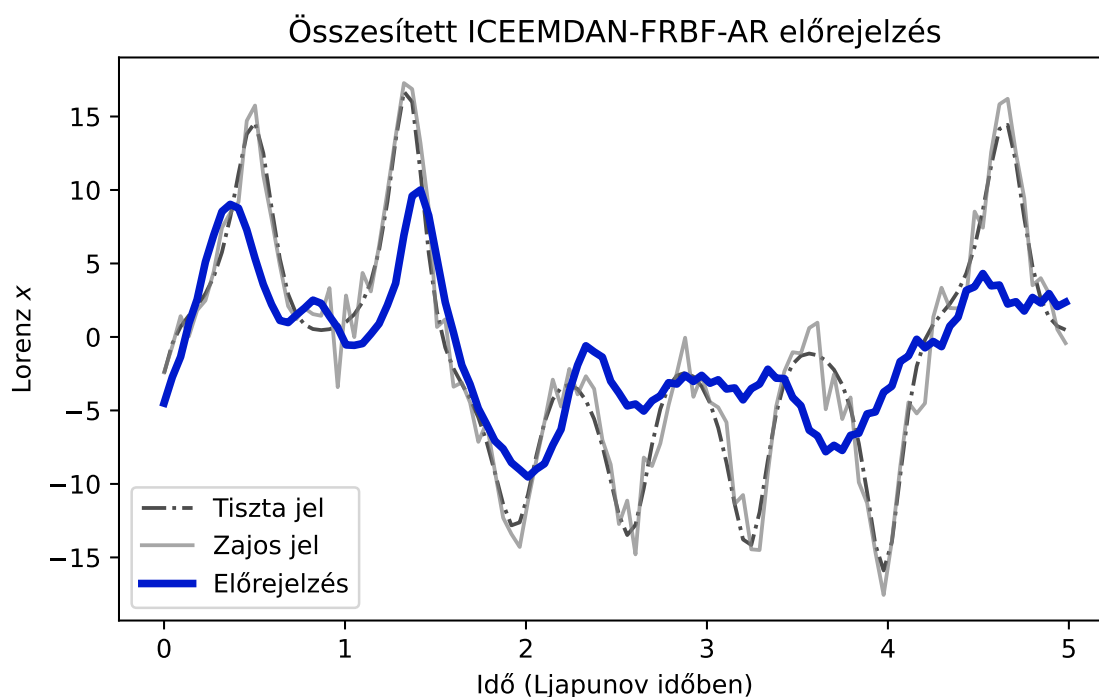
$$\text{MAAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \arctan \left(\frac{y_i - \hat{y}_i}{y_i} \right)$$

6.2.3. Normalizált Gyök Átlagos Négyzetes Eltérés

Szintén széleskörben elterjedt mérték, ami a négyzetes hibának egy változata, ami invariáns a szórásra, így könnyebb összehasonlítani különböző szórású idősorokat. Értéke a $[0; \infty[$ tartományba esik, minél kisebb, annál pontosabb az előrejelzés.

$$\text{NRMSE} = \sqrt{\frac{1}{n} \frac{\sum_{i=1}^n (y_i - \hat{y}_{ii})^2}{\frac{1}{n} \sum_{i=1}^n y_i^2}}$$

6.3. Az ICEEMDAN-FRBF-AR modell előrejelzése



6.1. ábra. Az ICEEMDAN-FRBF-AR modell végleges előrejelzése (az IMF-ek előrejelzéseinek összege).

Előrejelzési horizont (Ljapunov-időben)	NG-RC			NG-RC-Ch			ICEEMDAN-FRBF-AR		
	R ²	MAAPE	NRMSE	R ²	MAAPE	NRMSE	R ²	MAAPE	NRMSE
1	0,946	58,4%	0,204	0,924	65%	0,237	0,365	66%	1,089
2	-1,462	82,4%	1,567	-1,244	84,8%	1,532	0,636	60,8%	0,885
3	-1,034	78%	1,411	-0,929	78,7%	1,379	0,693	53,8%	0,835
4	-0,931	78,3%	1,432	-0,823	77,7%	1,392	0,614	57,3%	0,957
5	-0,756	84,1%	1,386	-0,994	85,4%	1,407	0,585	59%	1,09

6.2. ábra. Az előrejelzések eredményeinek összehasonlítása. A mutatókat az előrejelzett idősor és az eredeti (zajmentes) Lorenz x között mértük.

7. fejezet

Összefoglalás

A dolgozatban áttekintettük a káoszelmélet releváns fogalmait (1), megvizsgáltuk, hogy parciális megfigyelések esetén hogyan alkalmazhatjuk a Takens-beágyazást az attraktor rekonstrukciójához (2.1). Áttekintettük a szakirodalomban jelenleg fellelhető legjobb modelleket kaotikus rendszerek előrejelzésére (NG-RC 2.2.1, NG-RC-Ch 2.2.2). Konstatáltuk, hogy additív zaj megjelenése esetén erősen romlik ezen modellek előrejelző képessége.

Megvizsgáltuk a nemlineáris jelek felbontásához használt empirikus módú dekompozíció algoritmusát, valamint változatait (3), amely olyan egyszerűbb jelekre bontotta az idősorunkat, amiket külön-külön már jóval pontosabban előre tudunk jelezni.

Megismerkedtünk ezen komponensek predikciójához használt radiális bázisfüggvény hálózatokkal és variánsaikkal (4), amelyek jó stabilitással tudják ezeket a komponenseket előrejelezni.

Bevezettük az ICEEMDAN-FRBF-AR modellt, amely az ICEEMDAN algoritmus segítségével először csökkenti a jelen lévő zajt, majd pedig szintén ezen algoritmust alkalmazva additív komponensekre bontja a zajcsökkentett jelet. Ezeket a komponenseket beágyazzuk a Takens-beágyazás segítségével, majd előrejelezzük őket a fuzzy RBF-AR modellel. Kihasználva az ICEEMDAN azon tulajdonságát, hogy az eredő komponensek összege a jellel egyezik meg, az egyes komponensek előrejelzéseinek összege lesz az eredeti jel végleges előrejelzése. A modellünket a Lorenz-rendszer x változójának előrejelzésével értékeltük ki (mesterségesen additív zajt adva ahhoz), ami bizonyítottan kaotikus, így tökéletes arra, hogy ezen modellek összehasonlítására használjuk.

A modellünkről elmondható, hogy előrejelzései - bár rövid-távon nem olyan pontosak, mint az NG-RC-n alapuló modelleké - hosszabb távon (2-5 Ljapunov-idő) jelentősen jobb pontossággal rendelkeznek.

Mind a természetben, mind az ember által alkotott környezetben számtalan nemlineáris és kaotikus rendszerrel találkozhatunk, így a javasolt modellünk alkalmazásai szerteágazóak lehetnek (pl.: szívritmus analízise, pénzügyi idősorok előrejelzése).

További kutatás szükséges e területen a hiperparaméterek megválasztásának finomításához, valamint annak eldöntéséhez, hogy melyik RBF-középpontok kerüljenek egy jellemzőhöz a fuzzy RBF-AR modell jellemzőinek kialakításakor.

Ezúton köszönöm Armai Zsoltnak a dolgozathoz nyújtott szakmai segítségüket.

Irodalom

- [1] Christopher M Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995.
- [2] B.T. Bosworth és tsai. “Estimating Signal-to-Noise Ratio (SNR)”. *IEEE Journal of Oceanic Engineering* 33.4 (2008. okt.), 414–418. old. ISSN: 1558-1691. DOI: 10.1109/joe.2008.2001780. URL: <http://dx.doi.org/10.1109/JOE.2008.2001780>.
- [3] Liangyue Cao. “Practical method for determining the minimum embedding dimension of a scalar time series”. *Physica D: Nonlinear Phenomena* 110.1–2 (1997. dec.), 43–50. old. ISSN: 0167-2789. DOI: 10.1016/S0167-2789(97)00118-8. URL: [http://dx.doi.org/10.1016/S0167-2789\(97\)00118-8](http://dx.doi.org/10.1016/S0167-2789(97)00118-8).
- [4] E.S. Chng, S. Chen és B. Mulgrew. “Gradient radial basis function networks for nonlinear and nonstationary time series prediction”. *IEEE Transactions on Neural Networks* 7.1 (1996), 190–194. old. DOI: 10.1109/72.478403.
- [5] Jaesung Choi és Pilwon Kim. *Signal-noise separation using unsupervised reservoir computing*. 2024. DOI: 10.48550/ARXIV.2404.04870. URL: <https://arxiv.org/abs/2404.04870>.
- [6] Marcelo A. Colominas, Gastón Schlotthauer és María E. Torres. “Improved complete ensemble EMD: A suitable tool for biomedical signal processing”. *Biomedical Signal Processing and Control* 14 (2014. nov.), 19–29. old. ISSN: 1746-8094. DOI: 10.1016/j.bspc.2014.06.009. URL: <http://dx.doi.org/10.1016/j.bspc.2014.06.009>.
- [7] Daniel J. Gauthier és tsai. “Next generation reservoir computing”. *Nature Communications* 12.1 (2021. szept.). ISSN: 2041-1723. DOI: 10.1038/s41467-021-25801-2. URL: <http://dx.doi.org/10.1038/s41467-021-25801-2>.
- [8] J. M. González-Camacho és tsai. “Genome-enabled prediction of genetic values using radial basis function neural networks”. *Theoretical and Applied Genetics* 125.4 (2012. máj.), 759–771. old. ISSN: 1432-2242. DOI: 10.1007/s00122-012-1868-9. URL: <http://dx.doi.org/10.1007/s00122-012-1868-9>.
- [9] Norden E. Huang és tsai. “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454.1971 (1998. márc.), 903–995. old. ISSN: 1471-2946. DOI: 10.1098/rspa.1998.0193. URL: <http://dx.doi.org/10.1098/rspa.1998.0193>.
- [10] J.-S.R. Jang. “ANFIS: adaptive-network-based fuzzy inference system”. *IEEE Transactions on Systems, Man, and Cybernetics* 23.3 (1993), 665–685. old. ISSN: 0018-9472. DOI: 10.1109/21.256541. URL: <http://dx.doi.org/10.1109/21.256541>.
- [11] Sungil Kim és Heeyoung Kim. “A new metric of absolute percentage error for intermittent demand forecasts”. *International Journal of Forecasting* 32.3 (2016. júl.), 669–679. old. ISSN: 0169-2070. DOI: 10.1016/j.ijforecast.2015.12.003. URL: <http://dx.doi.org/10.1016/j.ijforecast.2015.12.003>.

- [12] Chris C. King. “Fractal and chaotic dynamics in nervous systems”. *Progress in Neurobiology* 36.4 (1991), 279–308. old. issn: 0301-0082. DOI: [https://doi.org/10.1016/0301-0082\(91\)90003-J](https://doi.org/10.1016/0301-0082(91)90003-J). URL: <https://www.sciencedirect.com/science/article/pii/030100829190003J>.
- [13] S. Lloyd. “Least squares quantization in PCM”. *IEEE Transactions on Information Theory* 28.2 (1982. márc.), 129–137. old. issn: 0018-9448. DOI: 10.1109/tit.1982.1056489. URL: <http://dx.doi.org/10.1109/TIT.1982.1056489>.
- [14] Wolfgang Maass, Thomas Natschläger és Henry Markram. “Real-time computing without stable states: A new framework for neural computation based on perturbations”. *Neural computation* 14.11 (2002), 2531–2560. old.
- [15] Maryam Masnadi-Shirazi és Shankar Subramaniam. “Attractor Ranked Radial Basis Function Network: A Nonparametric Forecasting Approach for Chaotic Dynamic Systems”. *Scientific Reports* 10.1 (2020. márc.). issn: 2045-2322. DOI: 10.1038/s41598-020-60606-1. URL: <http://dx.doi.org/10.1038/s41598-020-60606-1>.
- [16] Jooung Park és Irwin W. Sandberg. “Approximation and Radial-Basis-Function Networks”. *Neural Computation* 5.2 (1993. márc.), 305–316. old. issn: 1530-888X. DOI: 10.1162/neco.1993.5.2.305. URL: <http://dx.doi.org/10.1162/neco.1993.5.2.305>.
- [17] Irmantas Ratas és Kestutis Pyragas. “Application of next-generation reservoir computing for predicting chaotic systems from partial observations”. *Phys. Rev. E* 109 (6 2024. jún.), 64215. old. DOI: 10.1103/PhysRevE.109.064215. URL: <https://link.aps.org/doi/10.1103/PhysRevE.109.064215>.
- [18] B. Samardzic és B.M. Zlatkovic. “Analysis of spatial chaos appearance in cascade connected nonlinear electrical circuits”. *Chaos, Solitons & Fractals* 95.C (2017), 14–20. old. DOI: 10.1016/j.chaos.2016.12.0. URL: <https://ideas.repec.org/a/eee/chsofr/v95y2017icp14-20.html>.
- [19] Floris Takens. “The reconstruction theorem for endomorphisms”. *Bulletin of the Brazilian Mathematical Society* 33 (2002), 231–262. old.
- [20] Maria E. Torres és tsai. “A complete ensemble empirical mode decomposition with adaptive noise”. *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2011. máj. DOI: 10.1109/icassp.2011.5947265. URL: <http://dx.doi.org/10.1109/ICASSP.2011.5947265>.
- [21] Markus Vogl, Milena Kojić és Petar Mitić. “Dynamics of green and conventional bond markets: Evidence from the generalized chaos analysis”. *Physica A: Statistical Mechanics and its Applications* 633 (2024), 129397. old. issn: 0378-4371. DOI: <https://doi.org/10.1016/j.physa.2023.129397>. URL: <https://www.sciencedirect.com/science/article/pii/S0378437123009524>.
- [22] ZHAOHUA WU és NORDEN E. HUANG. “ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOISE-ASSISTED DATA ANALYSIS METHOD”. *Advances in Adaptive Data Analysis* 01.01 (2009. jan.), 1–41. old. issn: 1793-7175. DOI: 10.1142/s1793536909000047. URL: <http://dx.doi.org/10.1142/S1793536909000047>.
- [23] Jia-Rong Yeh, Jiann-Shing Shieh és Norden E Huang. “Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method”. *Advances in adaptive data analysis* 2.02 (2010), 135–156. old.