

Empirijske ortogonalne funkcije

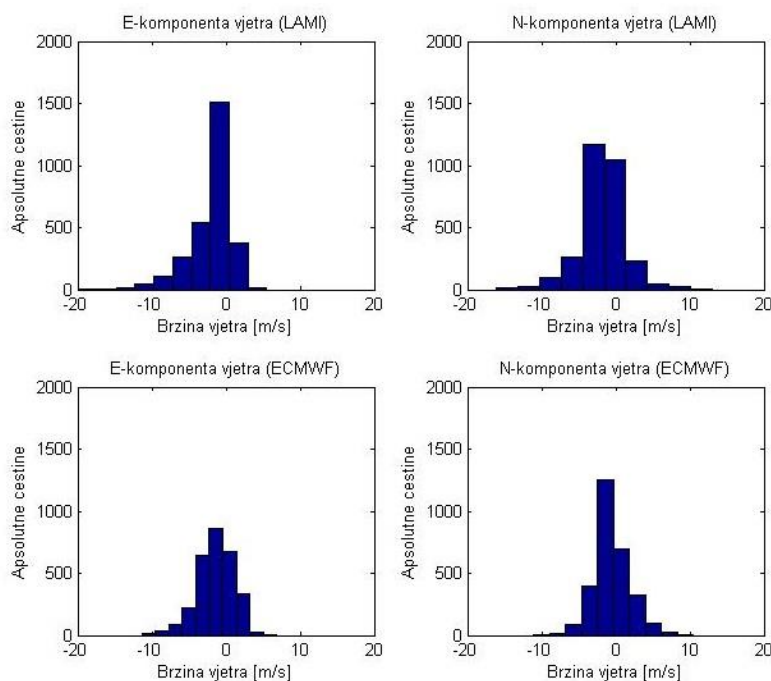
HELENA LATEČKI, II-134-19

Geofizički odsjek, PMF, Horvátovac ul. 95, 10 000 Zagreb

4.11.2021.

Za odabranu postaju Rijeka analizirani su trosatni vremenski nizovi temperature na 2 m iznad tla te E- i N-komponenta vjetra na 10 m visine. Nizovi potječu iz meteorološkog numeričkog modela LAMI (Italija) te modela Europskog centra za srednjoročnu prognozu vremena (ECMWF) te se odnose na razdoblje od 2.11.2002. do 30.9.2003.

Histogrami E- i N-komponente vjetra za oba numerička modela prikazani su na Slici 1. Najveće čestine postižu se za vrijednosti brzina bliskim nuli, a od dva dana modela, veći raspon brzina daje LAMI.



Slika 1. Histogrami trosatnih nizova E- i N-komponente vjetra za model LAMI i model ECMWF za razdoblje od 2. studenog 2002. do 30. rujna 2003. za postaju Rijeka.

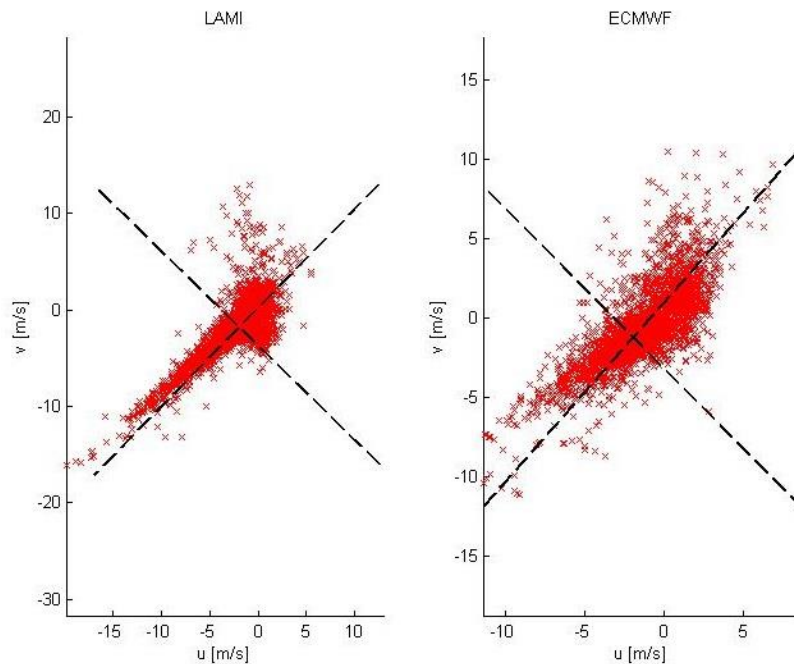
Da bi se odredile glavne osi vjetra, tj. smjer u kojem je on najjači te najslabiji, korištena je metoda empirijskih ortogonalnih funkcija (EOF). Cilj ove metode jest odrediti dominantne modove varijabilnosti stohastičkog polja, tj. preciznije odrediti obrasce simultane varijabilnosti. Za m -dimenzionani slučajni vektor $X(\omega) = (X_1(\omega), \dots, X_m(\omega))^T$ gdje je $X_i(\omega)$ slučajni vektor razmatranog stohastičkog polja (npr. temperature ili vjetra), i indeks postaje, ω vrijeme za koje ima n mjerenja, traži se rastav oblika:

$$X(\omega) = \alpha_1(\omega) \cdot e_1 + \dots + \alpha_m(\omega) \cdot e_m.$$

$\alpha_i(\omega)$ predstavlja vremenski niz (EOF koeficijente), a e_i prostorne modove (EOF-e) koji se određuju po principu najveće varijance, uz pretpostavku da je $\mathbb{E}(X) = 0$ kada je $\text{Var}(X) = \mathbb{E}(X^2)$. Traže se e_1 i $\alpha_1(\omega)$ tako da $X(\omega) - \alpha_1(\omega) \cdot e_1$ ima što manju varijancu, odnosno prostorni raspored koji kad se miče u vremenu skuplja što više varijance. Preostali prostorni modovi odabiru se po istom principu, tako da opišu što veći dio varijance ostatka.

Za zadni slučaj, primjenom ove metode moguće je odrediti glavne osi vjetra - $\alpha_1(\omega)$ bit će iznos, a e_1 smjer najveće brzine vjetra. Za provedbu EOF metode, oformljena je matrica X tipa $m \times n$ koja po stupcima sadrži mjerenja pojedine postaje, a po redcima vremenske nizove mjerenja E- i N-komponente vjetra. Nad takvom je matricom potom proveden SVD-rastav $X = USV^*$ jer je matrica $U = [u_1, \dots, u_m]$ zapravo matrica EOF modova (maksimum se postiže za $e_1 = u_1$), a SV^* su EOF koeficijenti. Stupci matrice U ujedno su i svojstveni vektori kovarijancijske matrice $Q = \frac{1}{n} X X^*$ i cilj je naći $\text{Var}(X(\omega), e_1) = e_1^* Q e_1 \rightarrow \max$.

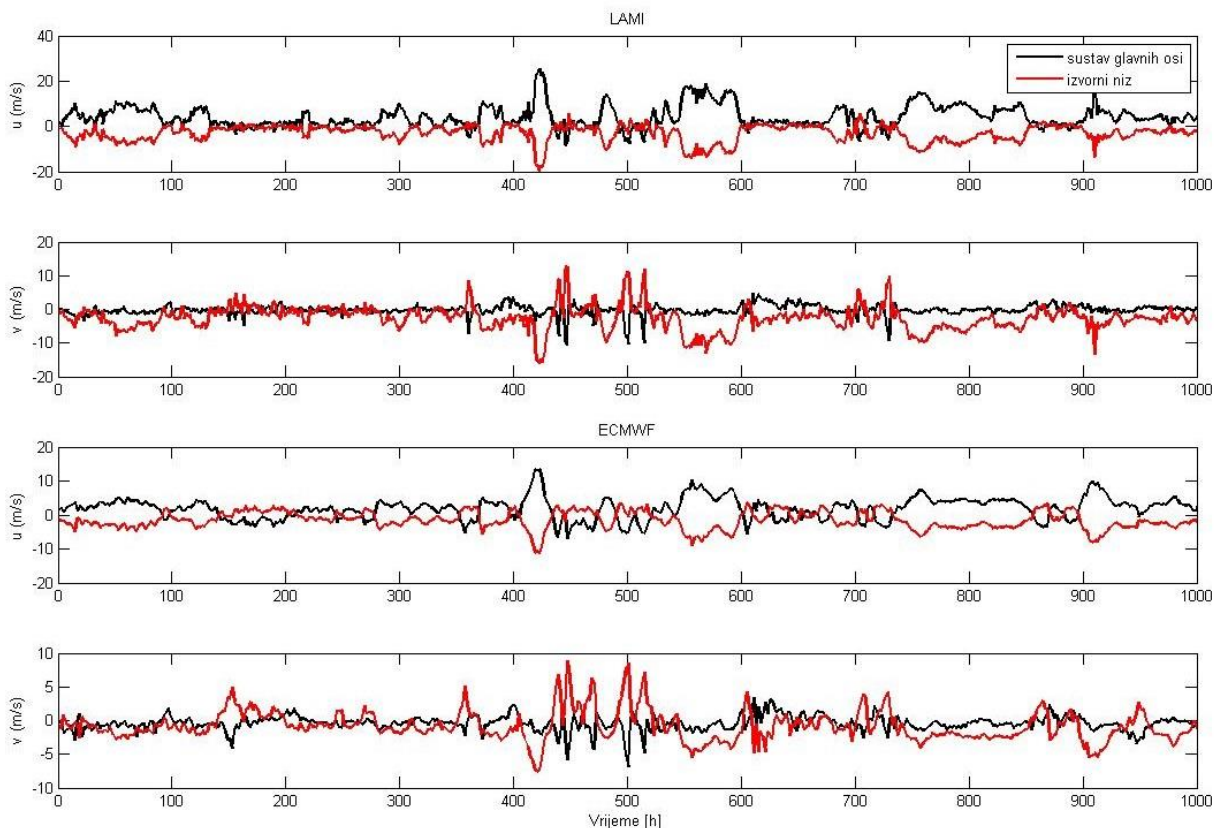
Dobiveni smjerovi glavnih osi iznose 135.59° i -134.41° za model LAMI, te 138.49° i -131.51° za model ECMWF. Na Slici 2. prikazani su dijagrami raspršenja i glavne osi vjetra.



Slika 2. Dijagrami raspršenja brzine vjetra za model LAMI i ECMWF. Crne crtane linije označavaju glavne osi vjetra.

Srednjkvadratne brzine komponenti vjetra u E- i N-smjeru iznose 12.75 m/s i 12.32 m/s za model LAMI, te 8.07 m/s i 6.69 m/s za model ECMWF. U sustavi glavnih osi, vrijednosti iznose 22.98 m/s i 2.09 m/s za model LAMI, te 13.10 m/s i 1.66 m/s za model ECMWF. Dakle, u sustavu glavnih osi, E-komponenta brzine vjetra veća je od N-komponente, što se moglo naslutiti na temelju Slike 2.

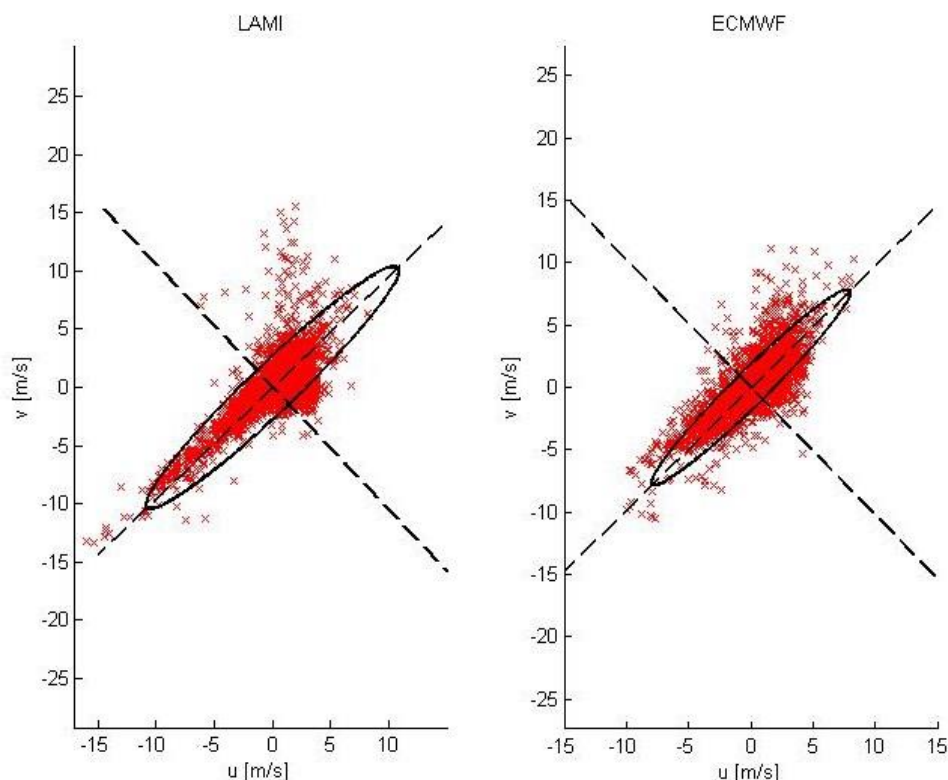
Na Slici 3. prikazane su komponente vjetra u sustavu glavnih osi zajedno s izvornim nizovima. U sustavu glavnih osi može se primijetiti povećanje E- i smanjenje N-komponente vjetra (što je jasno u skladu s iznosima srednjekvadratnih vrijednosti). Razlika dvaju vremenskih nizova (izvornog i onog u sustavu glavnih osi) izraženija je u slučaju modela LAMI. Bez obzira na razmatrani sustav i model, energija vjetra trebala bi ostati očuvana. Za zadani slučaj, razlike energija koje se dobe reda su veličine greške zaokruživanja, pa se stoga može zaključiti da energija zaista ostaje očuvana, tj. da ne ovisi o odabiru koordinatnog sustava.



Slika 3. Izvorni vremenski nizovi E- i N-komponente (u i v) vjetra te nizovi u sustavu glavnih osi za model LAMI i ECMWF.

Nadalje, određeni su smjerovi najveće i najmanje varijabilnosti vjetra, mjereni standardnim devijacijama. Postupak je isti kao i ranije, stoga da su na početku oduzete srednje vrijednosti vremenskih nizova. Samim time, zapravo se radi s kovarijancijskom matricom, a ne matricom drugih momenata oko nule (dobiveni vremenski nizovi sada su centrirani oko nule). Pomoću SVD-rastava određuju se EOF koeficijenti i EOF-i koji ukazuju na smjerove glavnih osi tzv. elipse standardnih devijacija, tj. smjerove najveće i najmanje varijabilnosti vjetra. Duljina poluosi elipse u smjeru pripadnog EOF-a dana je odgovarajućom svojstvenom vrijednošću kovarijancijske matrice, što je upravo i standardna devijacija u traženom smjeru.

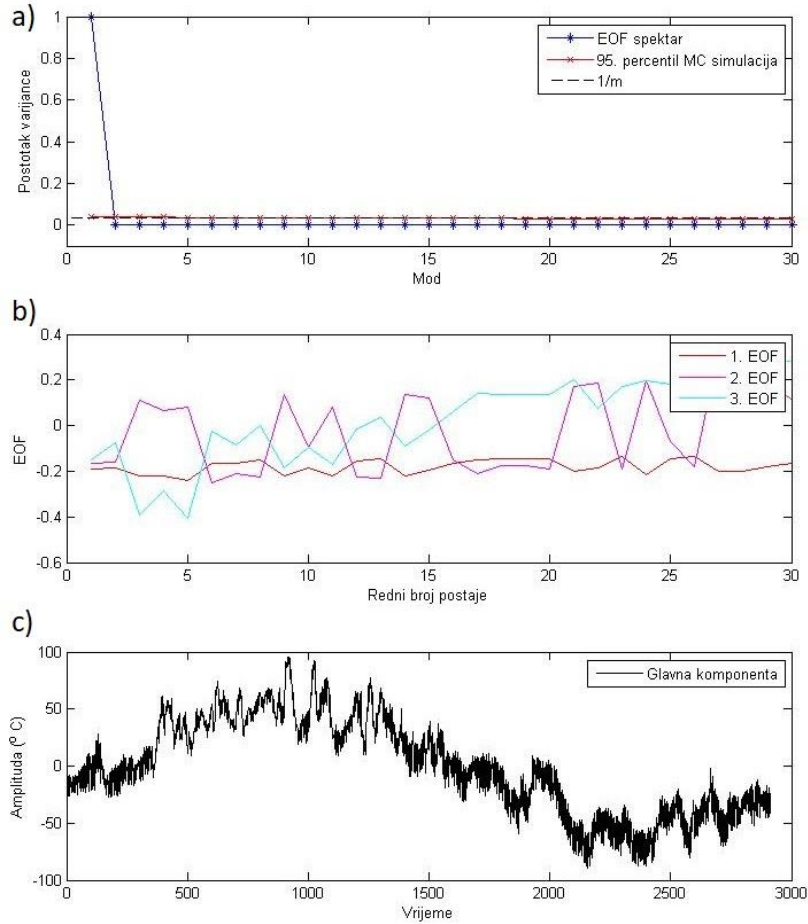
Na Slici 4. prikazani su dijagrami raspršenja te elipse standardnih devijacija dvaju razmatranih modela. Smjerovi glavnih osi iznose 133.64° i -136.36° za model LAMI, te 134.41° i -135.59° za model ECMWF. Duljine poluosi elipse su $a = 14.95$ i $b = 2.04$ za model LAMI, te $a = 11.13$ i $b = 1.38$ za model ECMWF. Središte elipse nalazi se u točki (0,0) što je i za očekivati jer su na početku bili oduzeti srednjaci.



Slika 4. Dijagrami raspršenja brzine vjetra za model LAMI i ECMWF te pripadne elipse standardnih devijacija. Crne crtkane linije označavaju glavne osi vjetra s oduzetom srednjom vrijednosti.

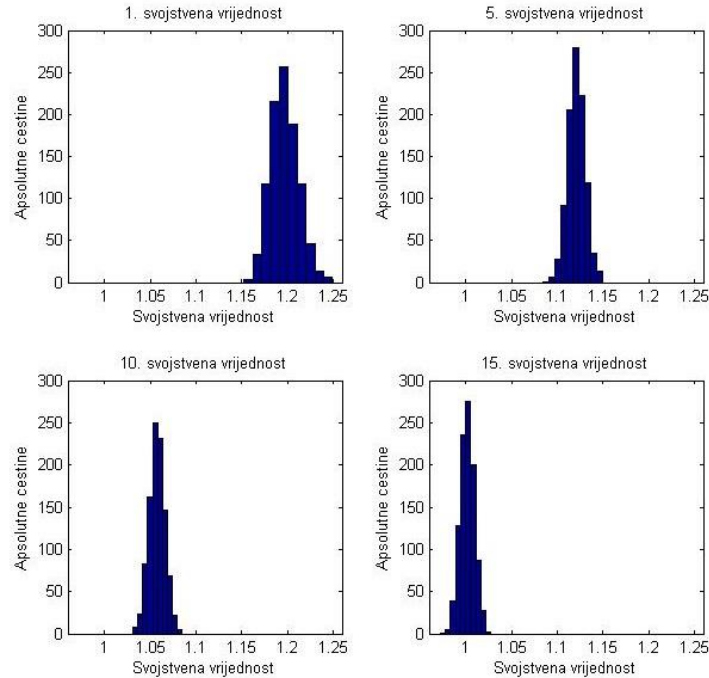
U nastavku je provedena EOF analiza za temperaturu iz LAMI modela. Na početku analize oduzeti su vremenski srednjaci. Kreirana je matrica X čiji su stupci indeksirani po postajama, a redci predstavljaju vremenske nizove temperature na odgovarajućoj postaji. EOF-i i EOF koeficijenti dobiveni su pomoću SVD-rastava.

Na Slici 5. prikazani su EOF spektar (svojstvene vrijednosti kovarijancijske matrice normirane na 100%) zajedno s 95. percentilom Monte Carlo (MC) simulacija i vrijednošću $1/m$ (gdje je m broj korištenih postaja), prva tri EOF-a i glavna komponenta (PC) prvog moda. S obzirom na „pravilo N“ (opisano u nastavku), značajan je samo prvi mod, tj. uzimanjem samo prvog moda u analizu moguće je opisati više od 95% varijance ulaznih podataka. Glavna komponenta je vremenski niz prvog EOF koeficijenta i govori o tome kako se prvi EOF mijenja u vremenu. Sa Slike 5. jasno se vidi da se prvi mod ponaša kao sinusoida čiji period iznosi otprilike godinu dana. To je i za očekivati s obzirom na to da je razmatran vremenski niz temperature.



Slika 5. a) EOF spektar, 95. percentil Monte Carlo (MC) simulacija i vrijednost $1/m$; b) Prva tri EOF-a; c) Glavna komponenta (PC) prvog moda.

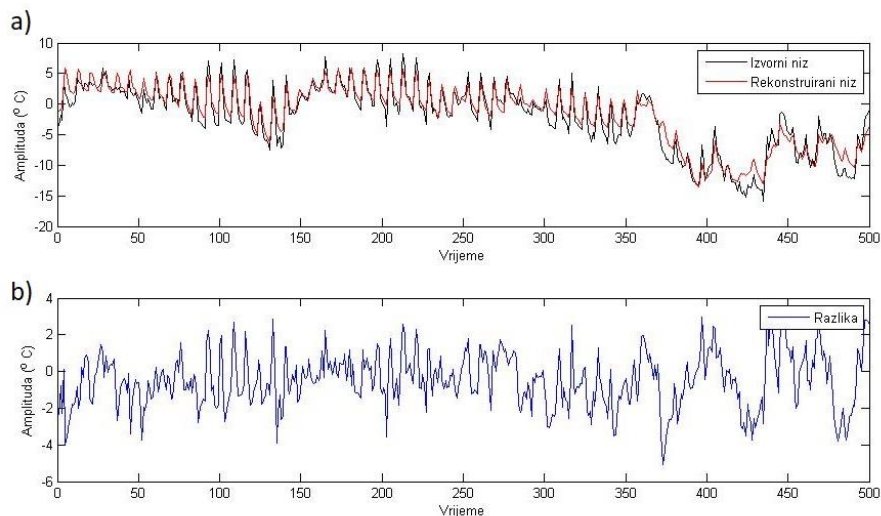
Pomoću „pravila N“ određeno je koji su EOF-i značajni. Ideja je da se 'fazni' prostor može podijeliti na dio koji sadrži fizikalni signal i dio koji sadrži šum. Pretpostavka je da signal nema malih svojstvenih vrijednosti te da se može opisati pomoću relativno malo modova. Granica za svojstvene vrijednosti λ_i kovarijancijske matrice ispod koje se nalazi isključivo šum određuje se tako da se pomoću MC metode simuliraju matrice z_i iz normalne razdiobe $z_i \sim N(0, \sigma^2 I)$. Dimenzije ovih matrica iste su kao i matrice X . Za svaku od njih odrede se svojstvene vrijednosti, normiraju, sortiraju te odrede pripadni 95. percentili realizacija pojedinih svojstvenih vrijednosti. Konkretno, pomoću Matlab funkcije *randn(m, n)* gdje je m broj postaja, a n duljina razmatranog vremenskog niza, sintetizirano je 1000 matrica prema normalnoj razdiobi. Za svaku od njih, izračunate su svojstvene vrijednosti te potom korištenjem Matlab funkcije *prctile()* određeni su 95. percentili. Sintetizirane matrice zapravo su matrice bijelog šuma i svojstvene vrijednosti koje su veće od 95. percentila smatraju se značajnima. Histogrami realizacija pojedinih svojstvenih vrijednosti sintetiziranih matrica prikazani su na Slici 6.



Slika 6. Histogrami realizacija prve, pete, desete i petnaeste normirane svojstvene vrijednosti matrica iz normalne razdiobe dobivenih MC metodom simulacija.

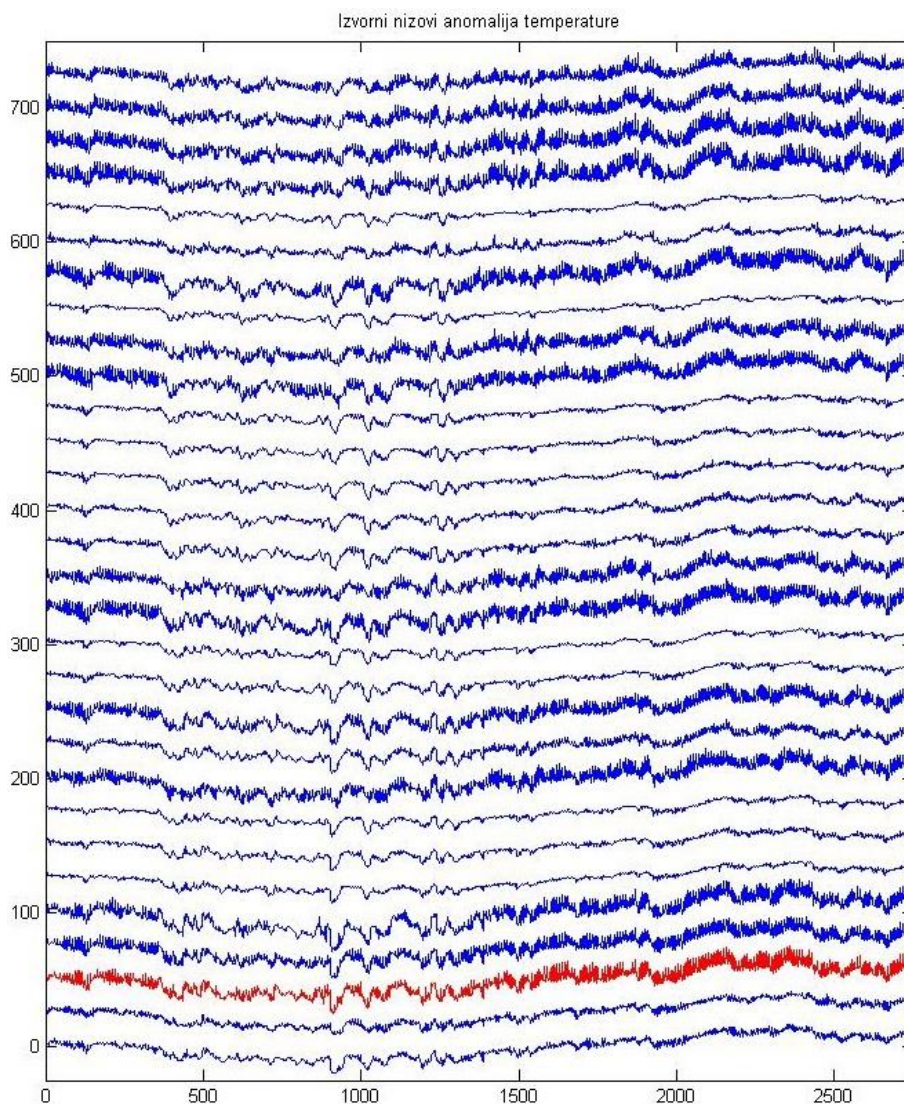
Histogrami zapravo predstavljaju realizacije procjenitelja svojstvenih vrijednosti za bijeli šum za pojedine postaje. Svojstvene vrijednosti u pravilu se gomilaju oko σ^2 . Za niz konačne duljine, očekivanja procjenitelja $E(\lambda_i)$ bit će veća od σ_i^2 za veće svojstvene vrijednosti, a manja za manje svojstvene vrijednosti.

S obzirom na rezultate dobivene analizom EOF spektra, izvorni niz temperature za postaju Rijeka rekonstruiran je uzimajući u obzir samo prvi mod (Slika 7). Rekonstruirani niz dobro opisuje varijabilnost izvornog niza, te na njihovoj razlici nema uočljivih trendova. Stoga je za zaključiti da je rekonstrukcija niza uzimajući u obzir samo prvi mod zadovoljavajuća.



▲ Slika 7. a) Izvorni niz temperature iz modela LAMI te rekonstruirani niz dobiven uzimajući u obzir samo prvi mod; b) Razlika izvornog i rekonstruiranog niza.

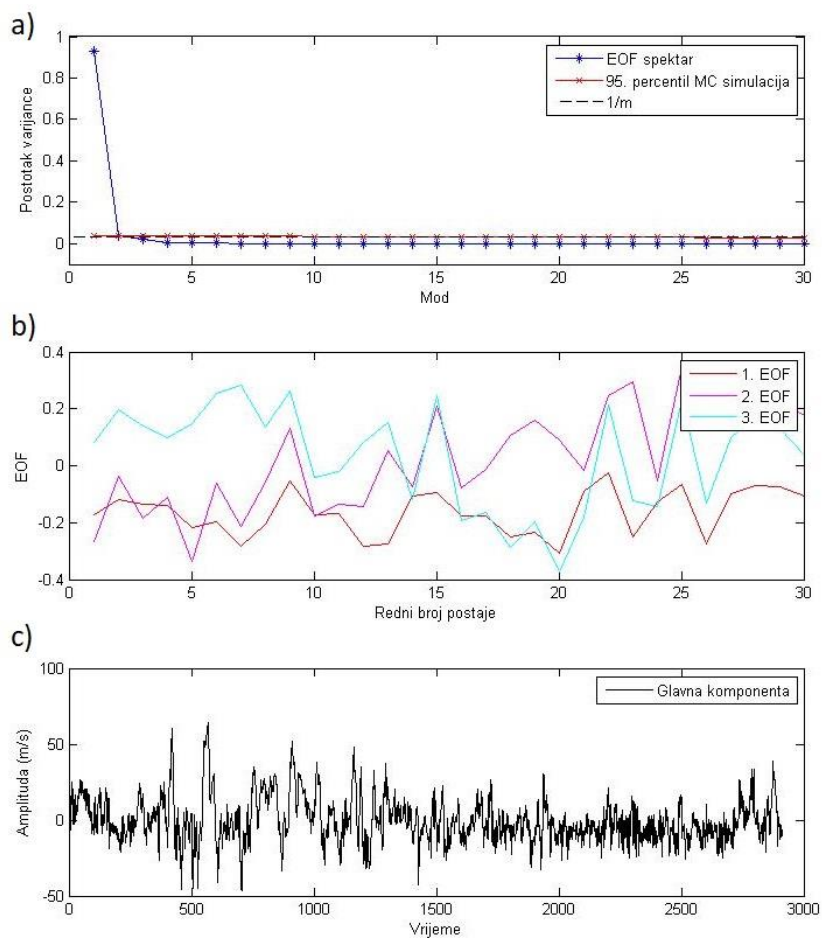
Na Slici 8. s posmakom su prikazani izvorni, centrirani nizovi (anomalije) svih postaja. Kao što je i u slučaju na Slici 5c, uočljiv je sinusoidalni oblik (s periodom od oko godine dana) u vremenu. Ponovno, usporedbom izvornih nizova i rezultata dobivenih za prvi EOF i glavnu komponentu, potvrđena je uspješnost rekonstrukcije glavnih obilježja izvornog niza uzimajući u obzir samo prvi mod.



Slika 8. Izvorni nizovi anomalija temperature svih postaja. Crvenom bojom naznačena je razmatrana postaja Rijeka.

Cijeli postupak potom je ponovljen za E-komponentu vjetra. Na Slici 9. prikazani su EOF spektar (svojtvene vrijednosti kovarijancijske matrice normirane na 100%) zajedno s 95. percentilom MC simulacija i vrijednošću 1/m, prva tri EOF-a i glavna komponenta (PC) prvog moda. S obzirom na „pravilo N“, uzimanjem prva dva moda moguće je opisati više od 95%

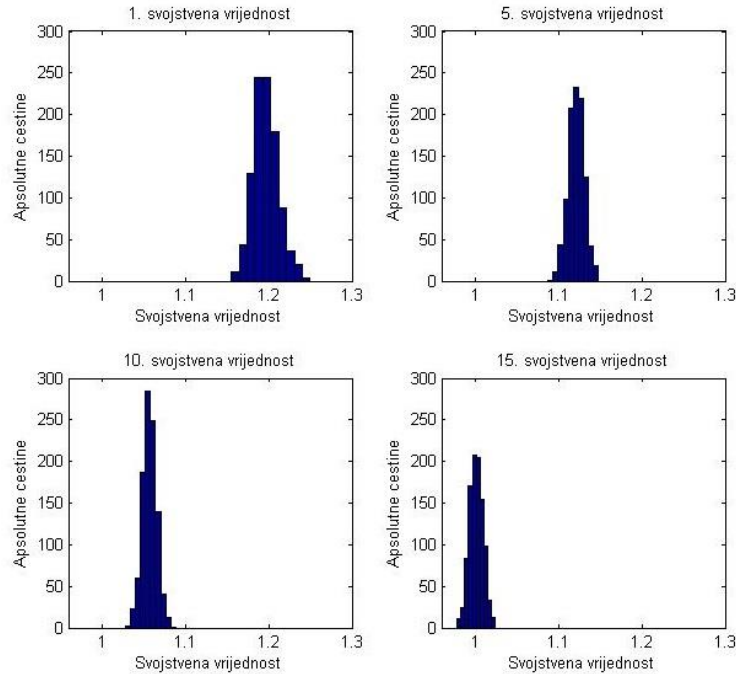
varijance ulaznih podataka. Na glavnoj komponenti nema uočljivog sinusoidalnog oblika kao što je u slučaju temperature, ali su varijabilnost i amplitude vremenskih nizova E-komponente vjetra izražene u prvom dijelu godine.



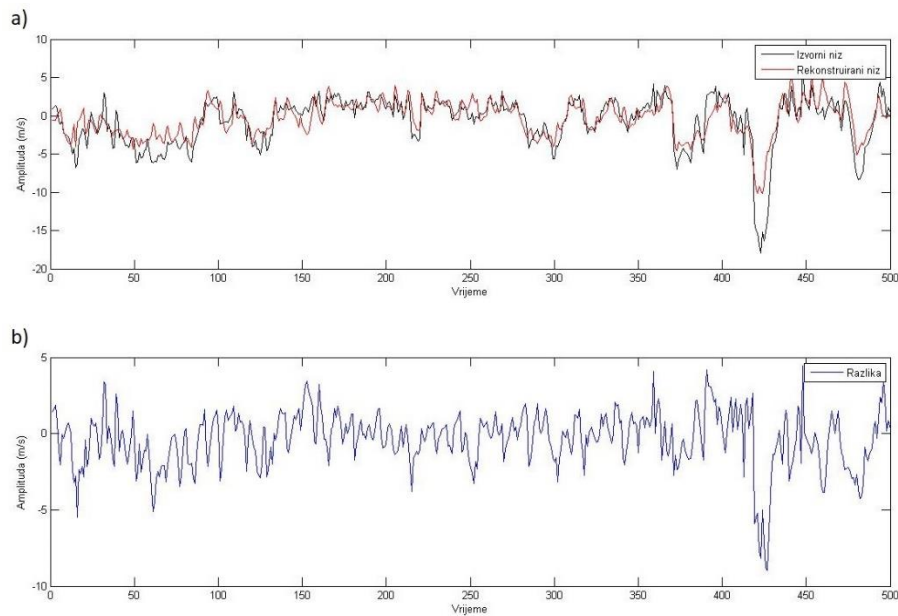
Slika 9. a) EOF spektar, 95. percentil Monte Carlo (MC) simulacija i vrijednost 1/m; b) Prva tri EOF-a; c) Glavna komponenta (PC) prvog moda.

Na histogramima realizacija pojedinih svojstvenih vrijednosti sintetiziranih matrica (Slika 10) ponovno se može uočiti grupiranje svojstvenih vrijednosti oko σ^2 .

Nadalje, s obzirom na rezultate dobivene analizom EOF spektra, izvorni niz E-komponente vjetra za postaju Rijeka rekonstruiran je uzimajući u obzir prva dva moda (Slika 11). Rekonstruirani niz relativno dobro opisuje varijabilnost izvornog niza. Razlika ovih dvaju nizova ne posjeduje ikakav specifičan trend, pa se može zaključiti da je rekonstrukcija niza uzimajući u obzir prva dva moda zadovoljavajuća.

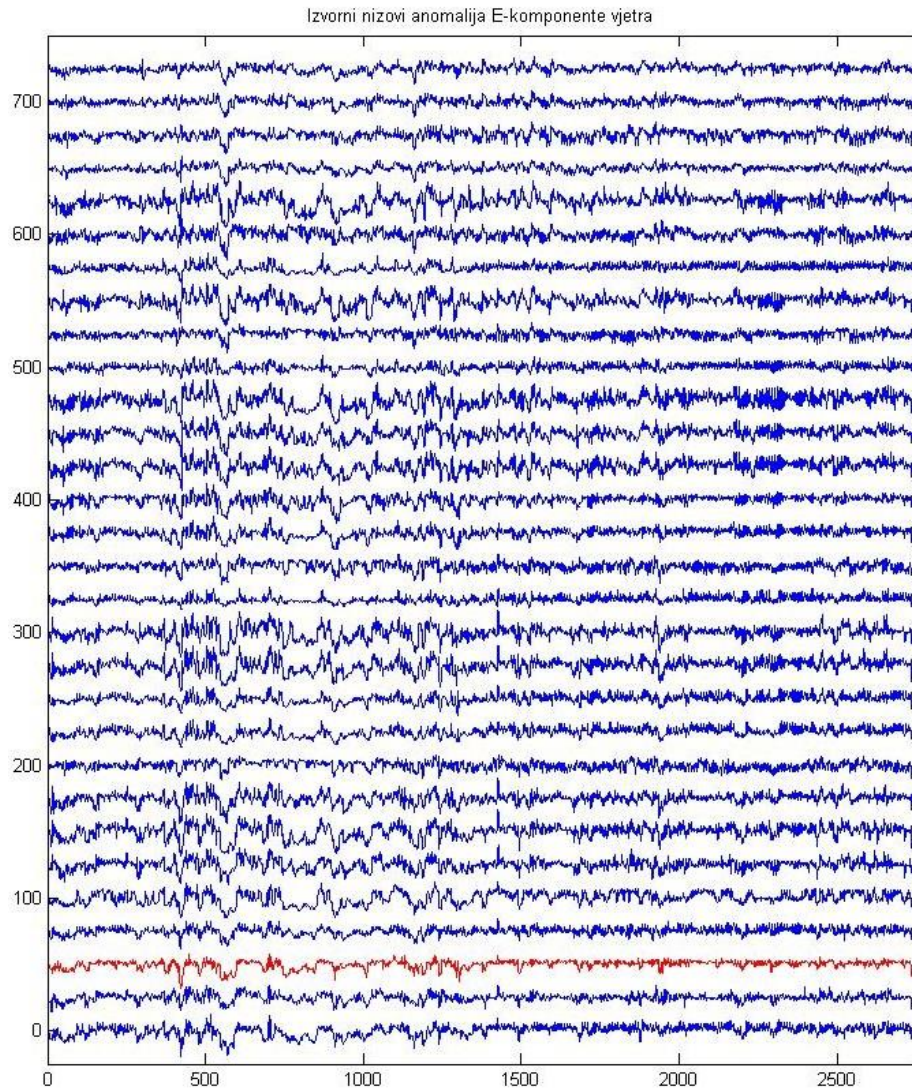


Slika 10. Histogrami realizacija prve, pete, desete i petnaeste normirane svojstvene vrijednosti matrica iz normalne razdiobe dobivenih MC metodom simulacija.



Slika 11. a) Izvorni niz E-komponente vjetra iz modela LAMI te rekonstruirani niz dobiven uzimajući u obzir samo prva dva moda; b) Razlika izvornog i rekonstruiranog niza.

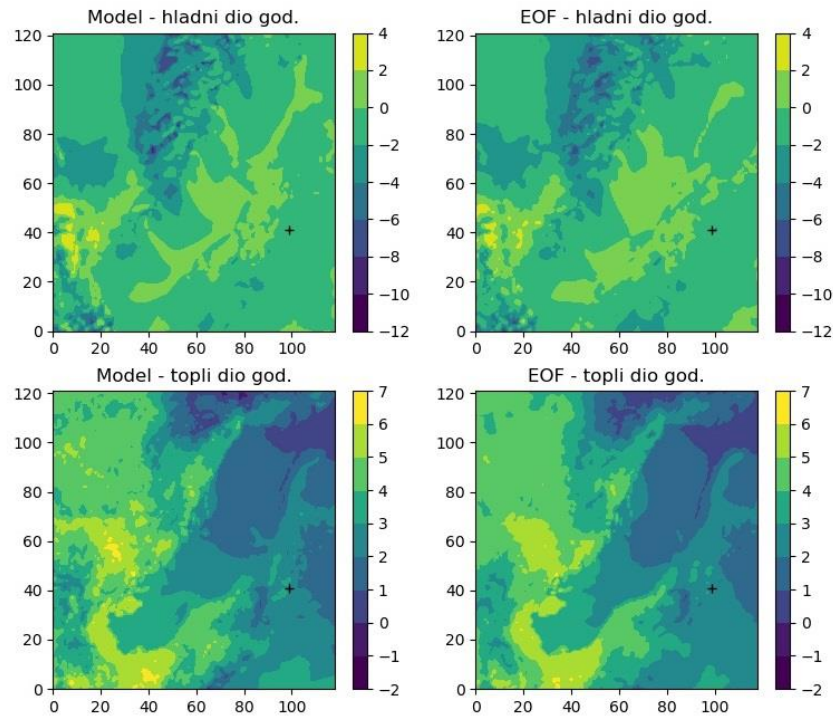
Na Slici 12. s posmakom su prikazani izvorni, centrirani nizovi (anomalije) svih postaja. Kao što je i u slučaju na Slici 9c, u prvom dijelu godine varijabilnost i amplitude vjetra su izraženije. Usporedbom izvornih nizova i glavne komponente, potvrđena je uspješnost rekonstrukcije glavnih obilježja izvornog niza.



Slika 12. Izvorni nizovi anomalija E-komponente vjetra svih postaja. Crvenom bojom naznačena je razmatrana postaja Rijeka.

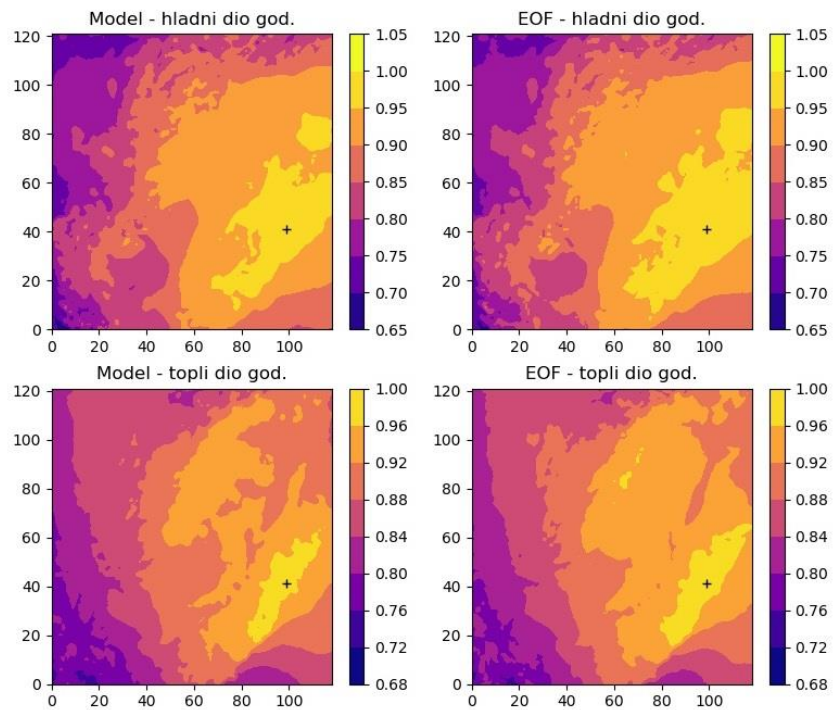
Za kraj, analizirane su korelacijske mape za dnevna polja temperature iz prognostičkog modela LAMI za razdoblje od 2.10.2005. do 1.10.2006. Korelacijska mapa za neko geofizičko polje (tj. za vremenski niz takvih polja) je slika (izolinije) koeficijenata korelacije između vrijednosti polja u zadanoj fiksnoj točki i svakoj drugoj točki istog polja.

Niz polja podijeljen je na dva dijela, hladni i topli dio godine, te su za svaku pojedinu postaju oduzete srednje vrijednosti. Nad poljem anomalija temperature provedena je EOF analiza te je polje anomalija rekonstruirano uzimanjem u obzir prvih 40 modova. Na Slici 13. prikazani su izvorni niz iz modela te rekonstruirani niz za 100-ti (hladni) dan te 300-ti (topli) dan. Rekonstruirani nizovi se relativno dobro slažu i opisuju prostornu varijabilnost izvornih nizova.



Slika 13. Izvorno (iz modela) i rekonstruirano (korištenjem EOF metode) polje anomalija temperature za hladni i topli dio godine. Crni križić predstavlja lokaciju točke (42, 100) koja je odabrana za daljnji račun korelacija.

Za fiksnu točku (42, 100) izračunata je i potom nacrtana korelacijska mapa za hladni i topli dio godine za izvorno polje iz modela te rekonstruirano polje dobiveno korištenjem EOF metode (Slika 14). Korelacijska mapa dobivena je tako da su izračunati koeficijenti korelacije između odabrane točke i svih ostalih točaka ('postaja'), pa stoga i ne čudi da su vrijednosti koeficijenata veće u blizini same odabrane točke. Korelacijske mape izvornog te rekonstruiranog polja relativno se dobro slažu jedna s drugom. Nadalje, koeficijenti rekonstruiranog polja prostorno na širem području poprimaju veće vrijednosti nego što je u slučaju izvornog polja. Izvorno polje u potpunosti opisuje signal, dok rekonstruirano velik dio varijance. Odabirom određenog broja EOF modova na neki se način vrši filtriranje niza, pa samim time rekonstruirano polje sadrži manje šuma od izvornog te su dobiveni koeficijenti korelacije veći. Sam račun korelacijske mape bit će brži uzimanjem manjeg broja EOF modova jer su matrice s kojima se provodi račun manjih dimenzija.



Slika 14. Korelacijska mapa za izvorno (iz modela) i rekonstruirano (korištenjem EOF metode) polje anomalija temperature za hladni i topli dio godine. Crni križić predstavlja lokaciju točke (42, 100) koja je odabrana za račun korelacija.