

Struktura mreže proteinskih interakcija i njena računarska analiza na primeru ispitivanja neuređenosti proteina u interaktomima virusa *SARS-CoV-2*

Lazar Vasović, 2006/2021

prof. dr Gordana Pavlović-Lažetić

Istraživanje podataka u bioinformatiči

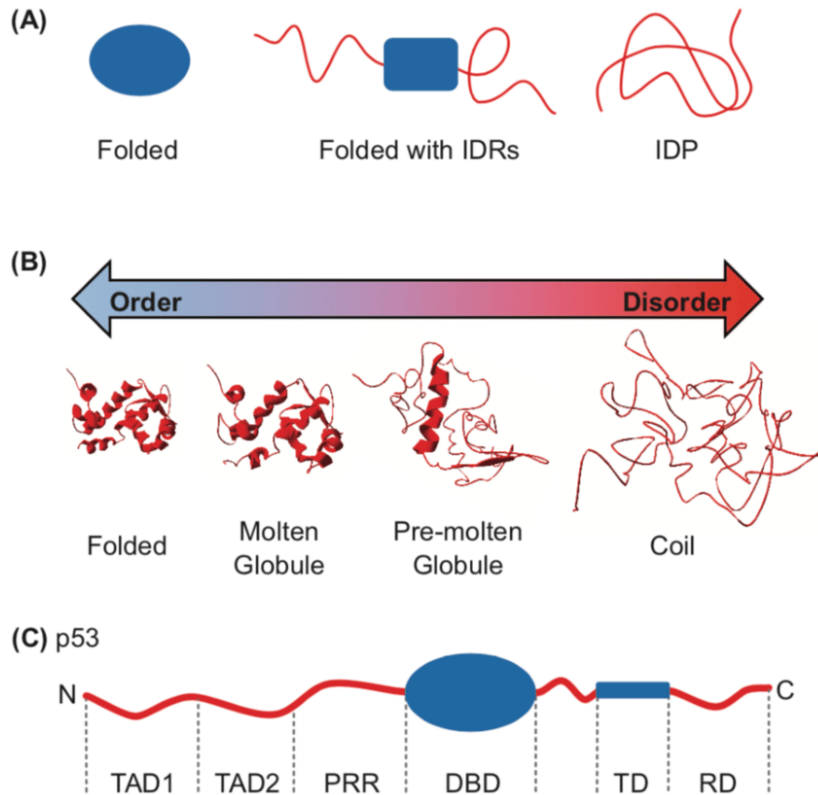
Seminar Katedre, 12. maj 2022.

Mreže proteinskih interakcija

- Biološki sistemi – interagujući blokovi
- Mreža (graf) – prirodni matematički model
- Interaktomi – mreže proteinskih interakcija
- Neusmerene grane



Neuređenost proteina



- Neuređeni proteini (*IDP*) – fleksibilna struktura
- Neuređeni regioni (*IDR*) – vrlo često sponje
- Neuređeni proteini → veliki interaktomi

Ovaj i povezani radovi

- Spajanje dve teme – poslednjih godina
- Eukariotske mreže – $ID(\text{čvorišta}) > ID(\text{ostali})$
- Virusne i kombinovane mreže – 🤔
- Osobine od interesa – stepen povezanosti i stepen neuređenosti prema nekoliko mera
- Uzajamno dejstvo – spojeni interaktom virusa (razmatrani *SARS-CoV-2*) i domaćina (čovjek)

Odabrani interaktomi

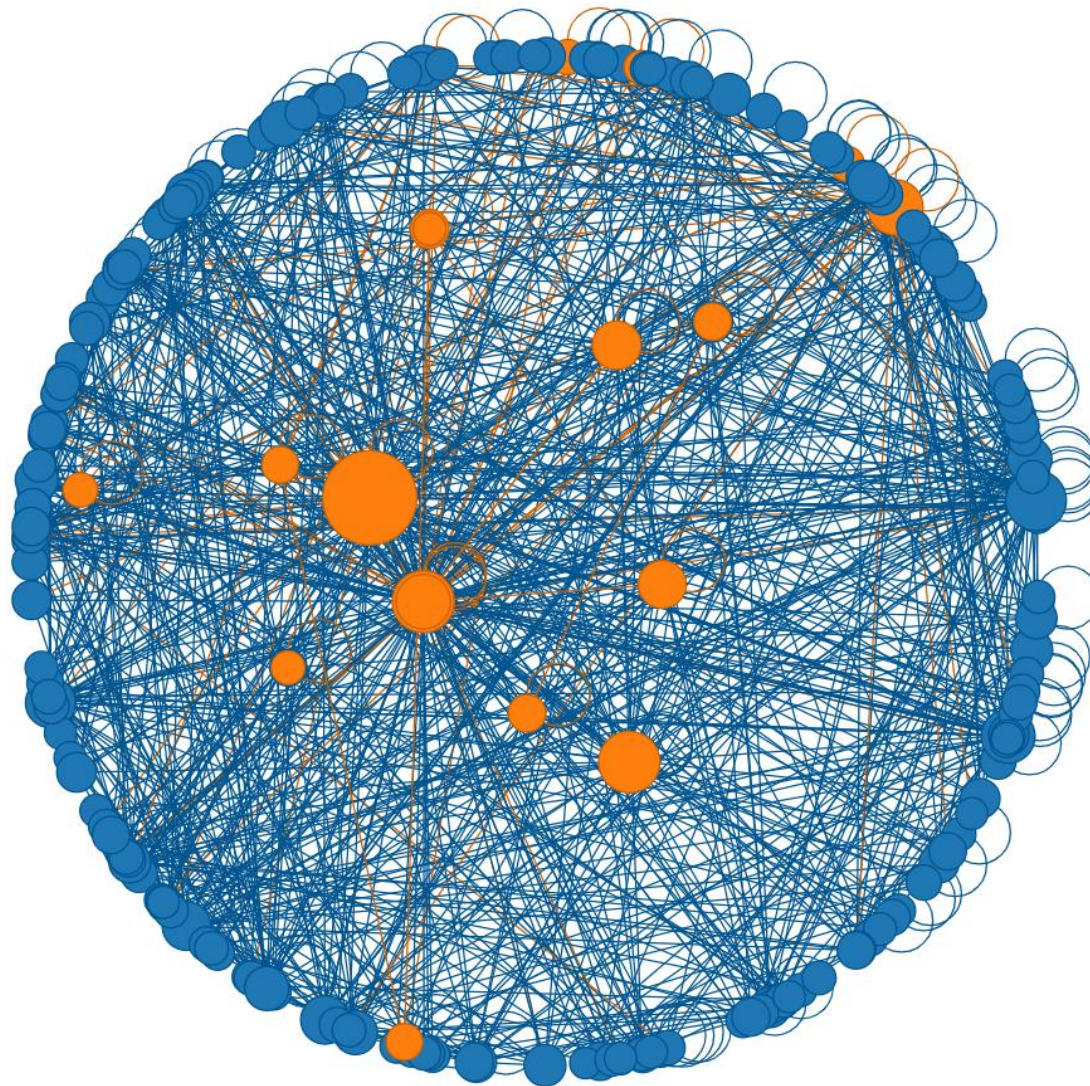
- Javno dostupni podaci – baze *IntAct*, *iRefIndex*

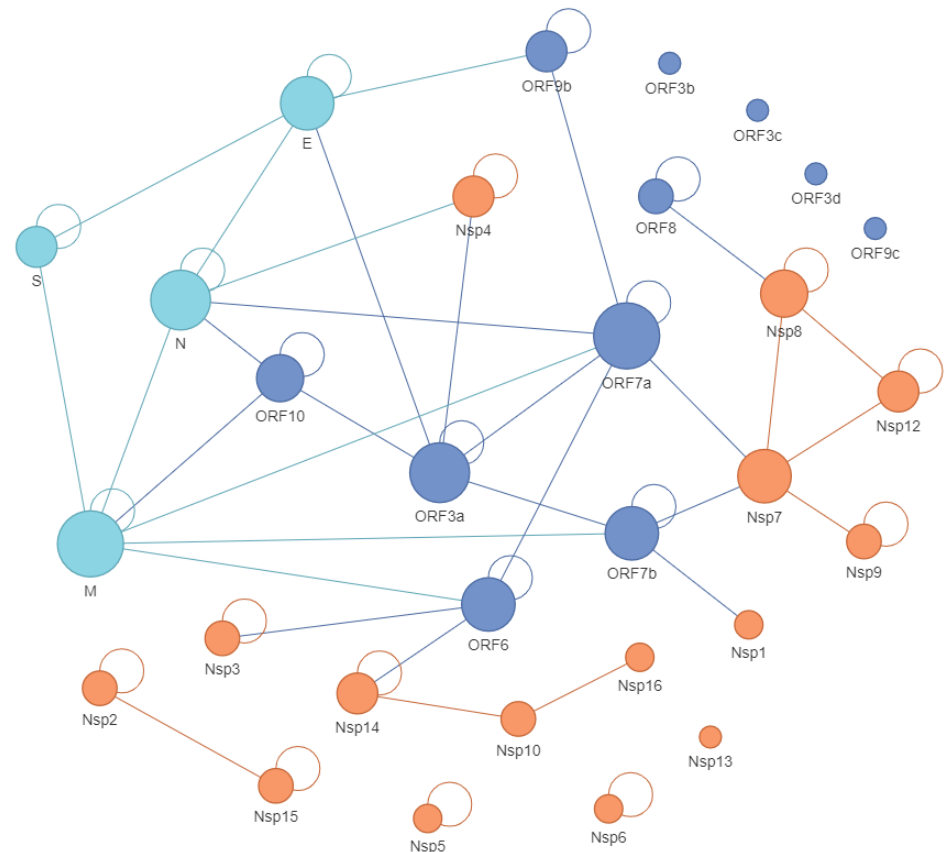
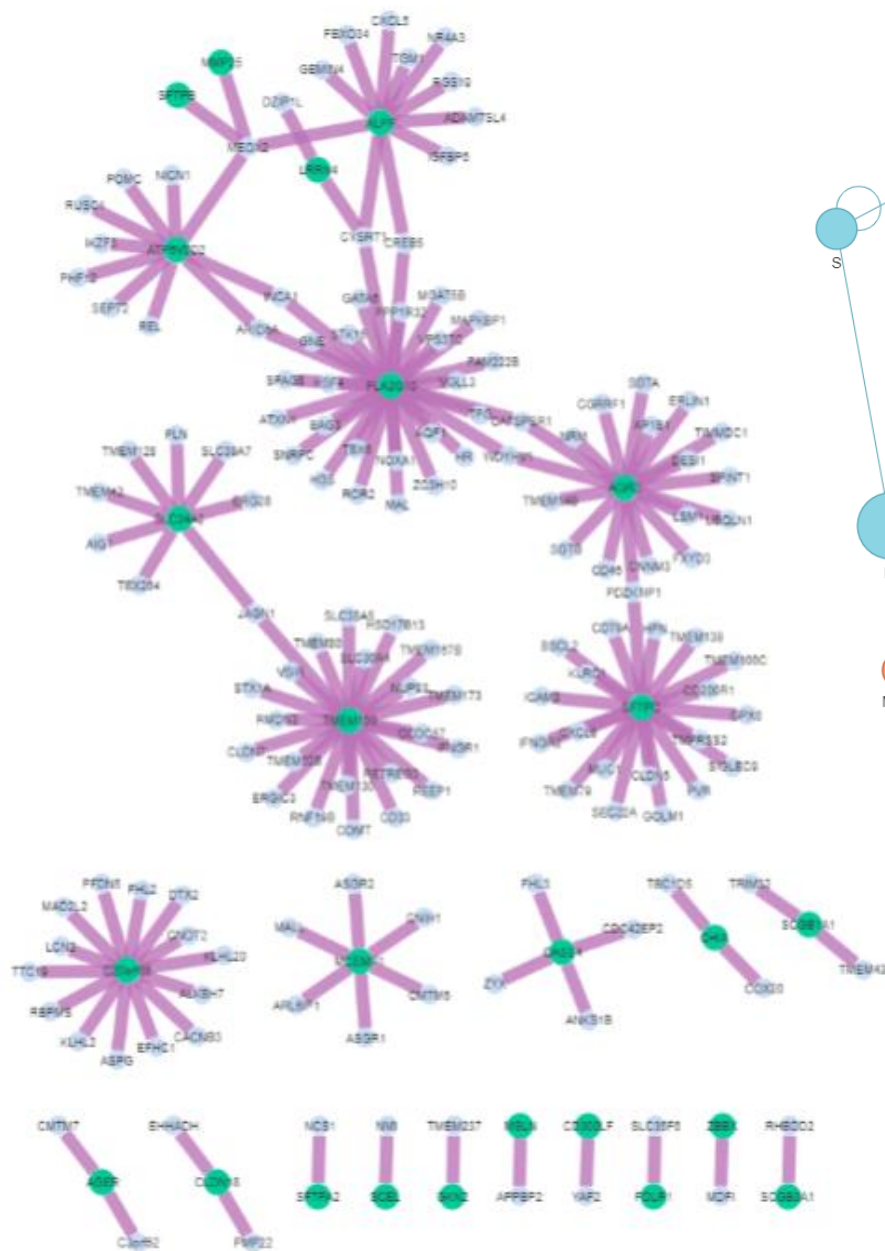
	Glavni proteini	Izvorna baza	Čvorovi	Grane
M_IntAct	membranski	IntAct	259	568
M_iRef	membranski	iRefIndex	211	486
Nsp_iRef	nestrukturni	iRefIndex	311	1122
SARS2_iRef	ceo proteom	iRefIndex	1419	8942

- Uzajamno dejstvo – mreža *HuRI*, baza *NDEx*

	Glavni proteini	Izvorna baza	Čvorovi	Grane
SARS_iRef	samo virusni	iRefIndex	30	53
pluća_HuRI	plućno tkivo	HuRI/NDEx	175	162

Primer interaktoma

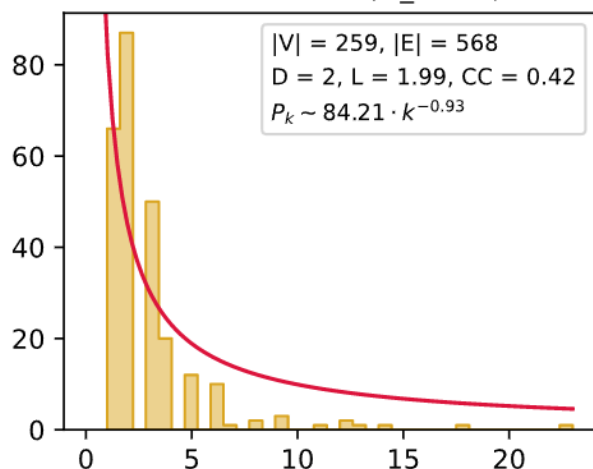




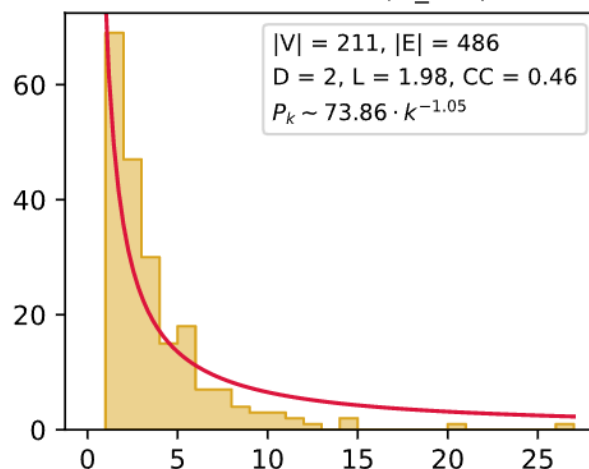
- Grafički prikaz – oznake, veličina, boja, tip čvora...

Raspodela stepena čvorova

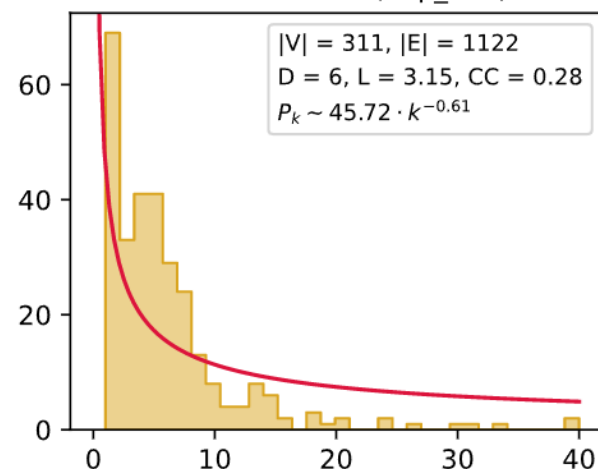
Osobine mreže (M_IntAct)



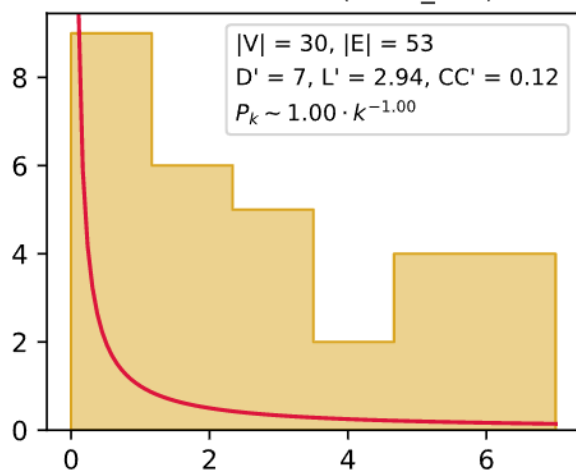
Osobine mreže (M_iRef)



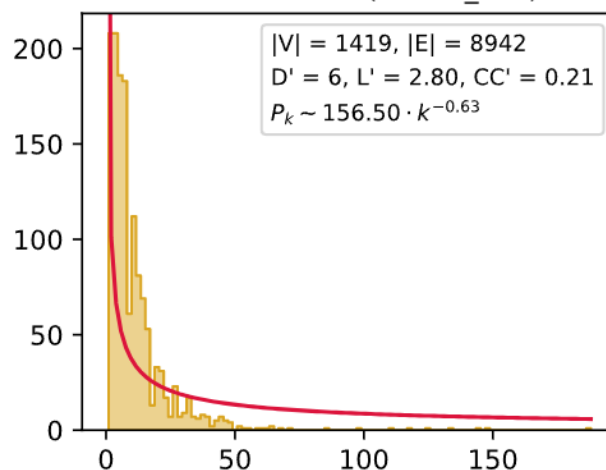
Osobine mreže (Nsp_iRef)



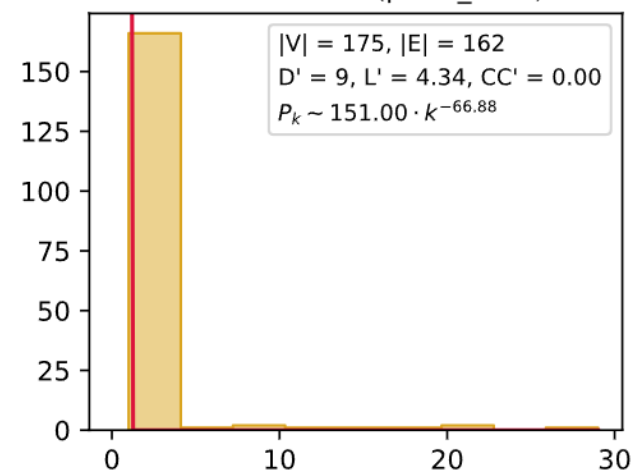
Osobine mreže (SARS_iRef)





Osobine mreže (SARS2_iRef)



Osobine mreže (pluća_HuRI)

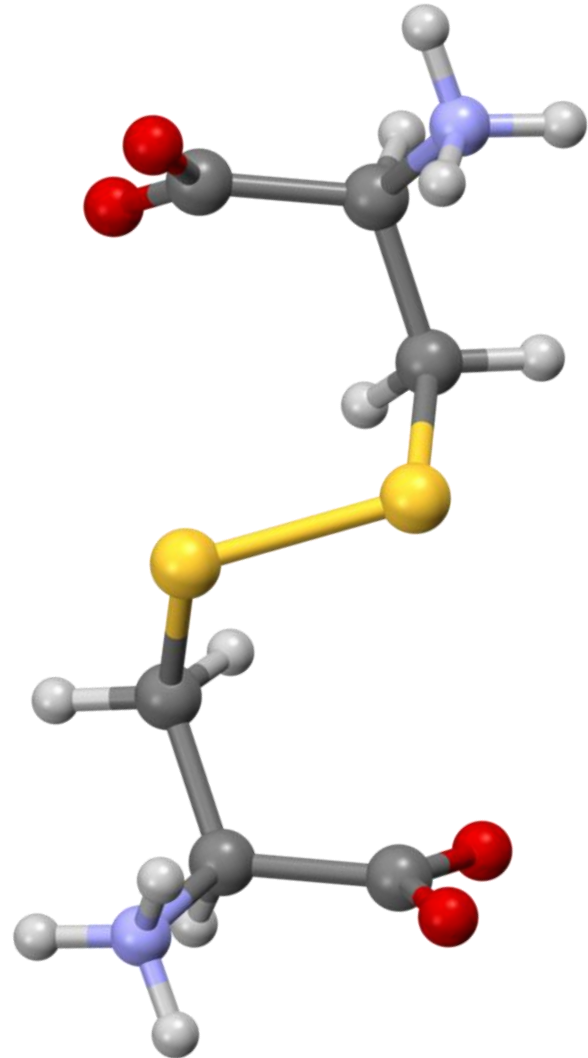
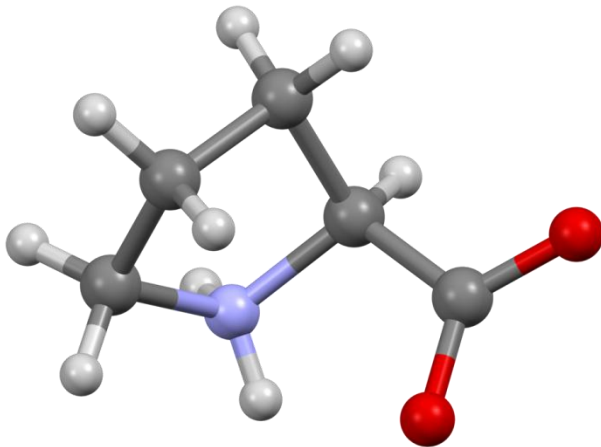


Mere neuređenosti

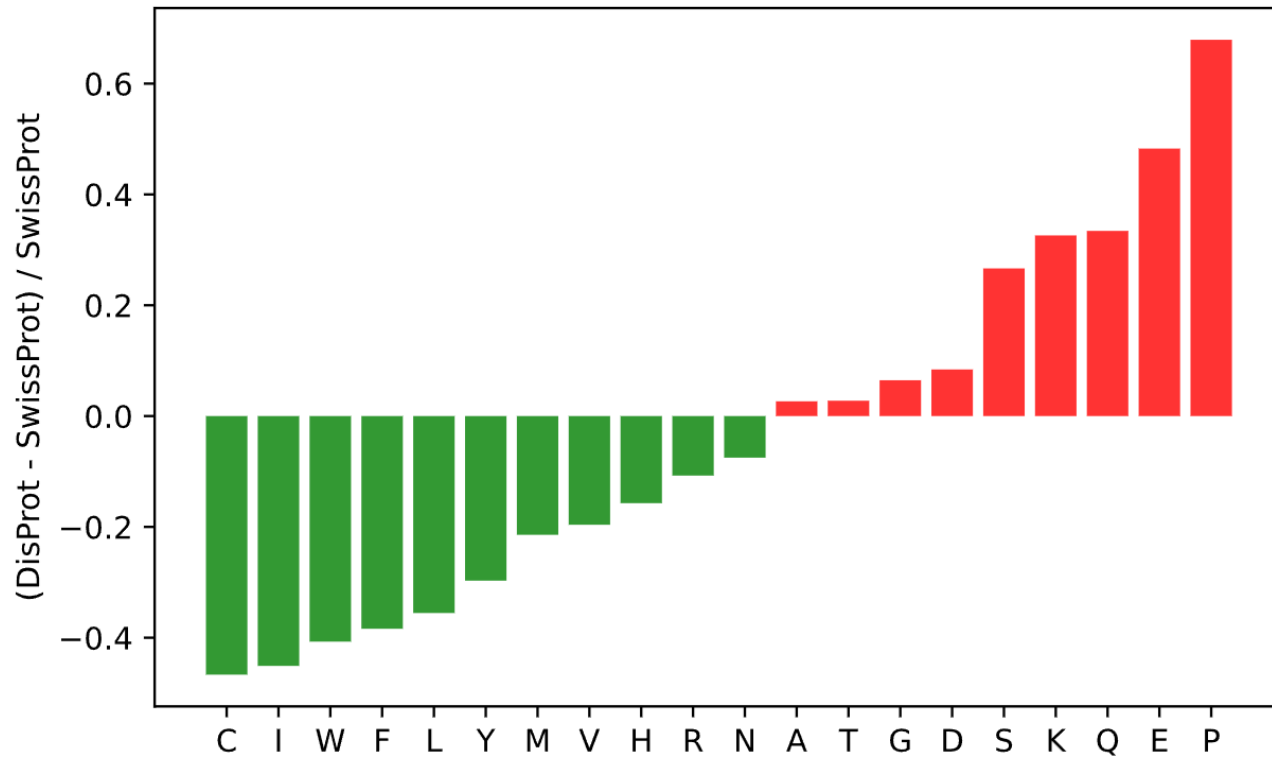
- Specijalizovane baze – *DisProt* 
- Alternativa – profilisanje i predviđanje
- Profilisanje – aminokiselinski sastav
- Predviđanje – ansambli neuronskih mreža
- Dodatne mere – neuređeni regioni, dužina
- Druga strana interakcije – susedi 

Aminokiselininski profil

- Statističko opravdanje – zastupljenost u bazama
- Biofizičko opravdanje – naelektrisanje, polarnost, veličina bočnog ostatka...

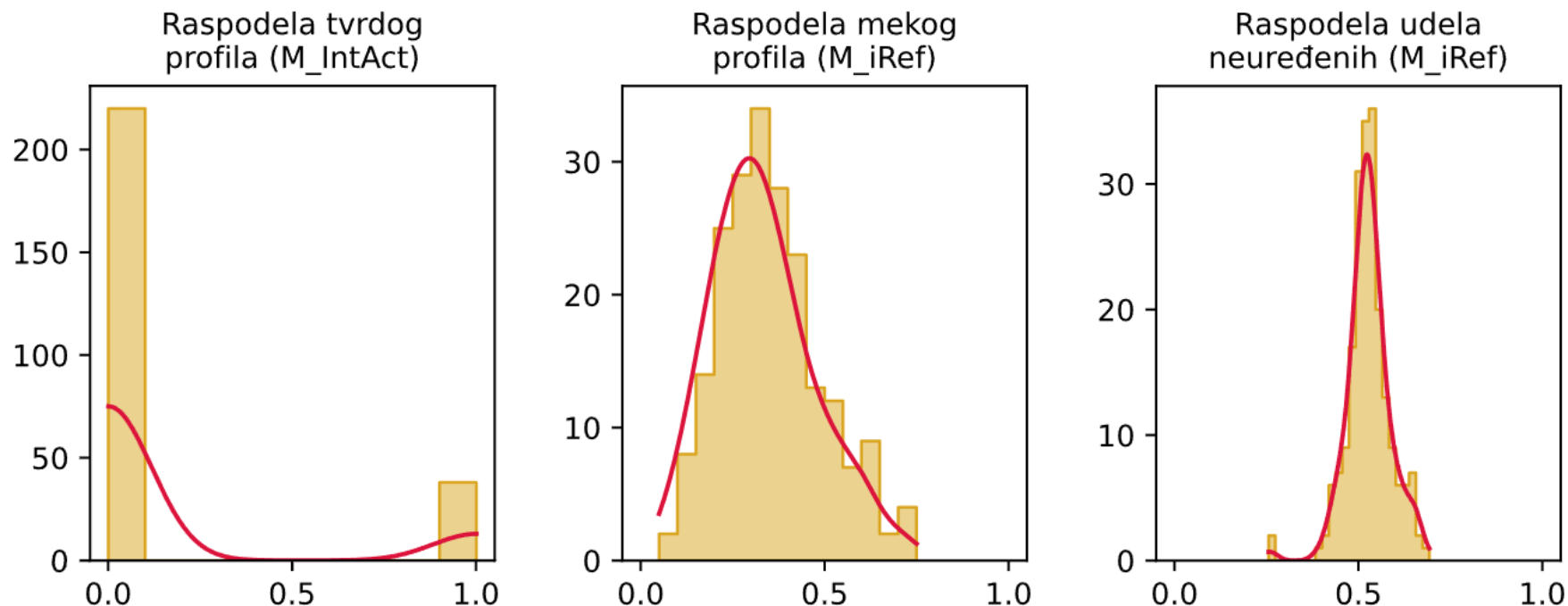


Razlike u udelima aminokiselina



- *Composition Profiler*, <http://www.cprofiler.org/help.html>

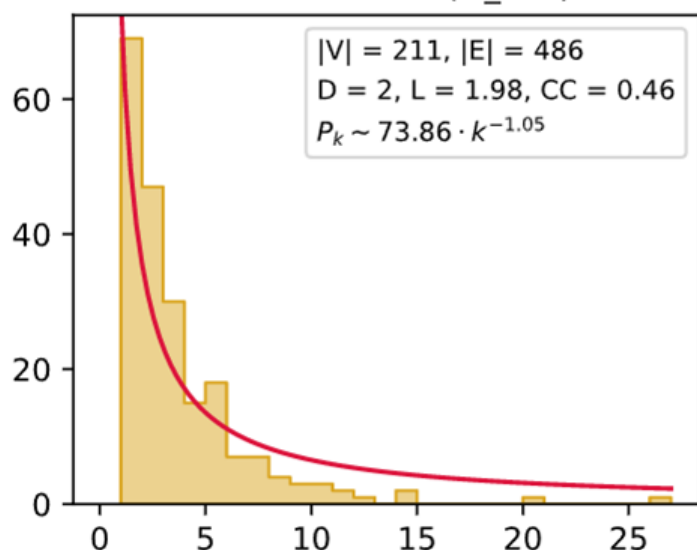
Tipovi aminokiselinskih profila



- Udeo „neuređenih” – prosta zastupljenost
- „Tvrdi” profil – kojoj grupi je bliži u proseku
- „Meki” profil – otežani prosek bliskosti

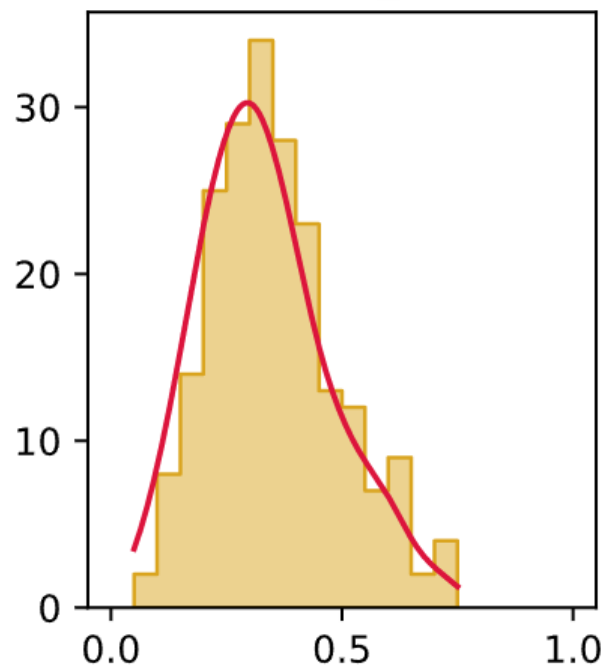
Jedna dimenzija → dve dimezije

Osobine mreže (M_iRef)



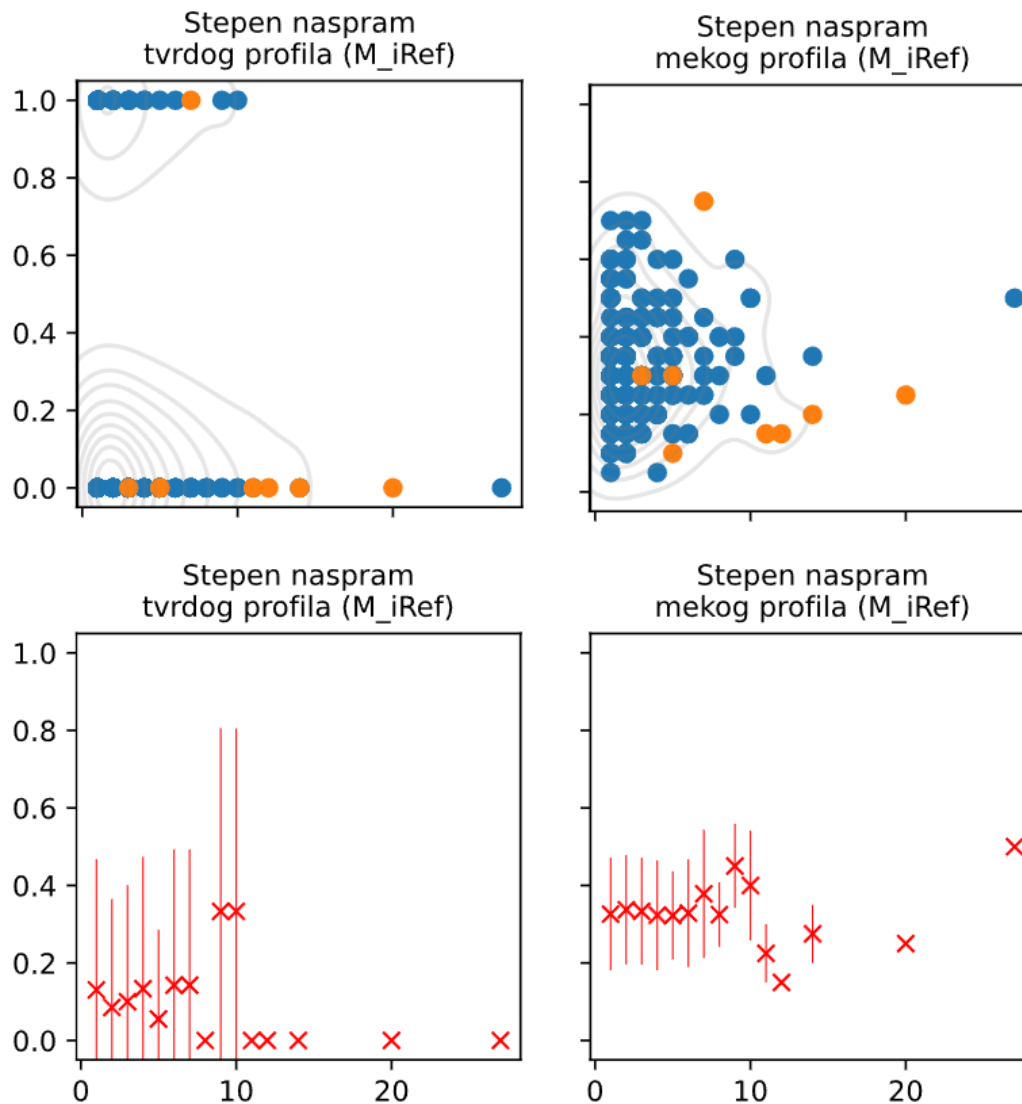
+

Raspodela mekog
profila (M_iRef)

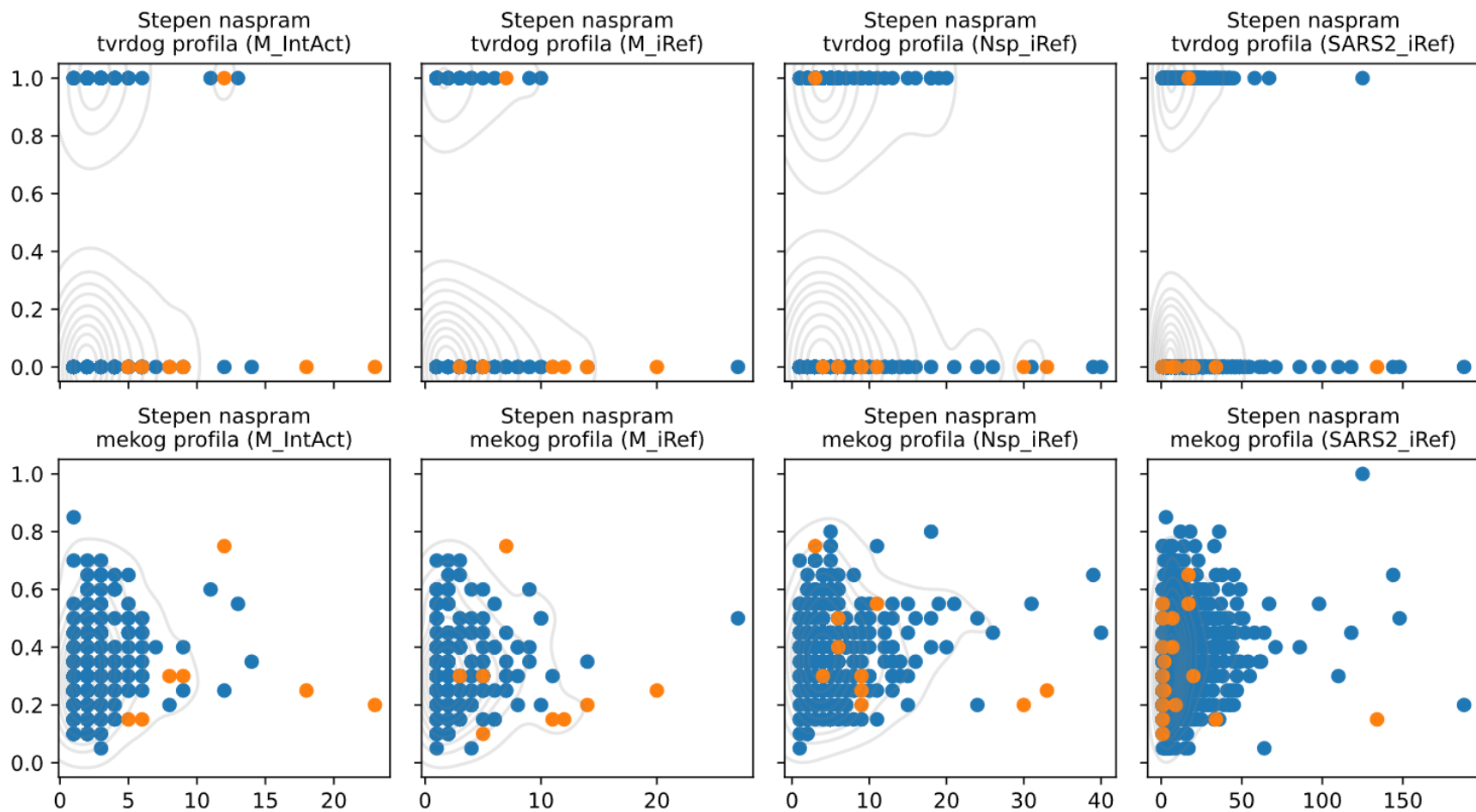


= ?

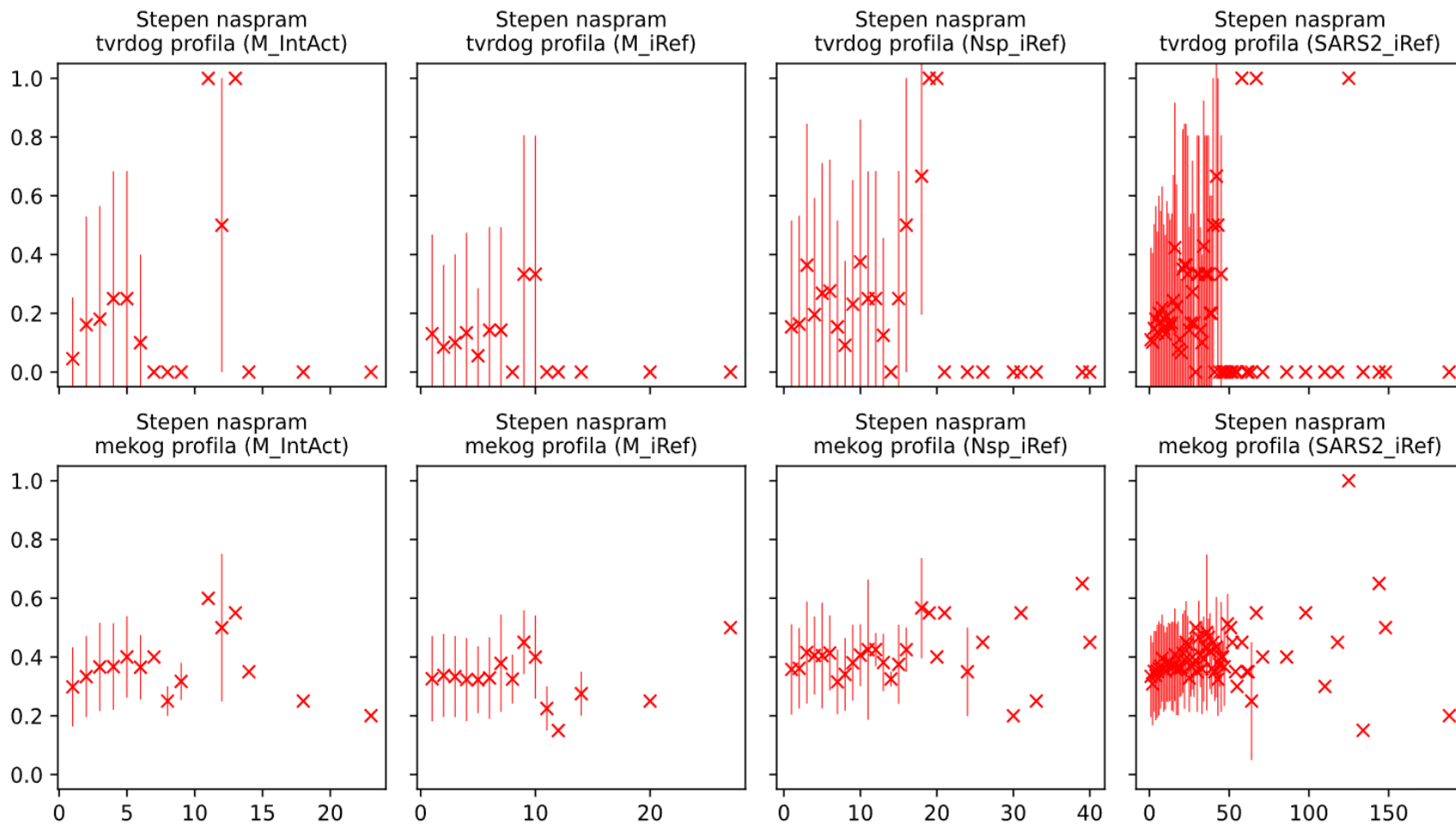
Zajednička i uslovna raspodela



Trend zajedničke raspodele



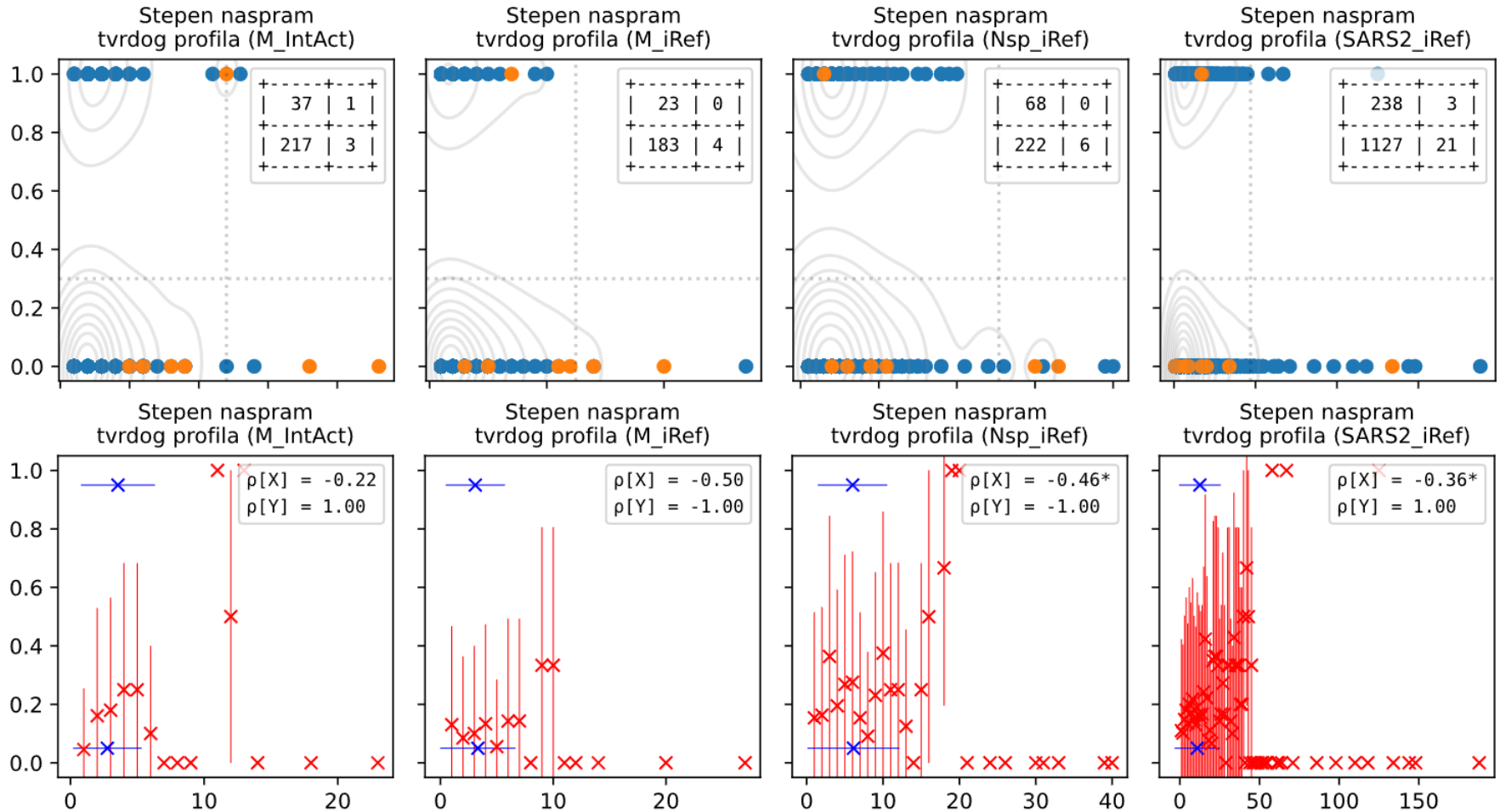
Trend uslovne raspodele



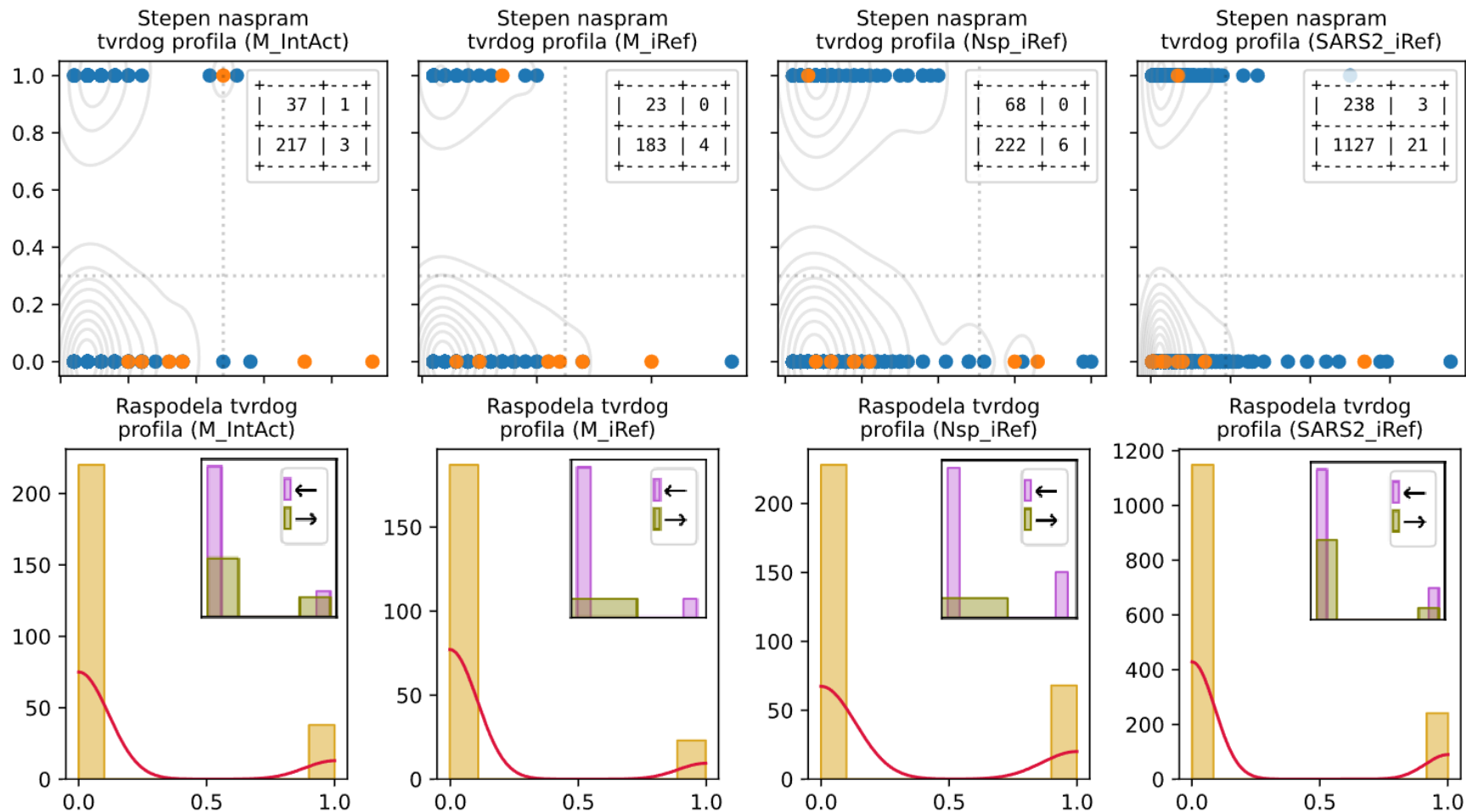
Granice podele

- Podela po horizontalnoj osi:
 - Precizni maksimumi – 5/14, 10, 21, 47
 - Aproksimacija sva četiri – kvantil 0,982
 - Približne vrednosti – 12, 12, 25, 47 ✓
- Podela po vertikalnoj osi:
 - Standard iz literature – 0,3
 - „Tvrđi” profil – uklapa se ✓
 - Zanimljivost – blizu medijane

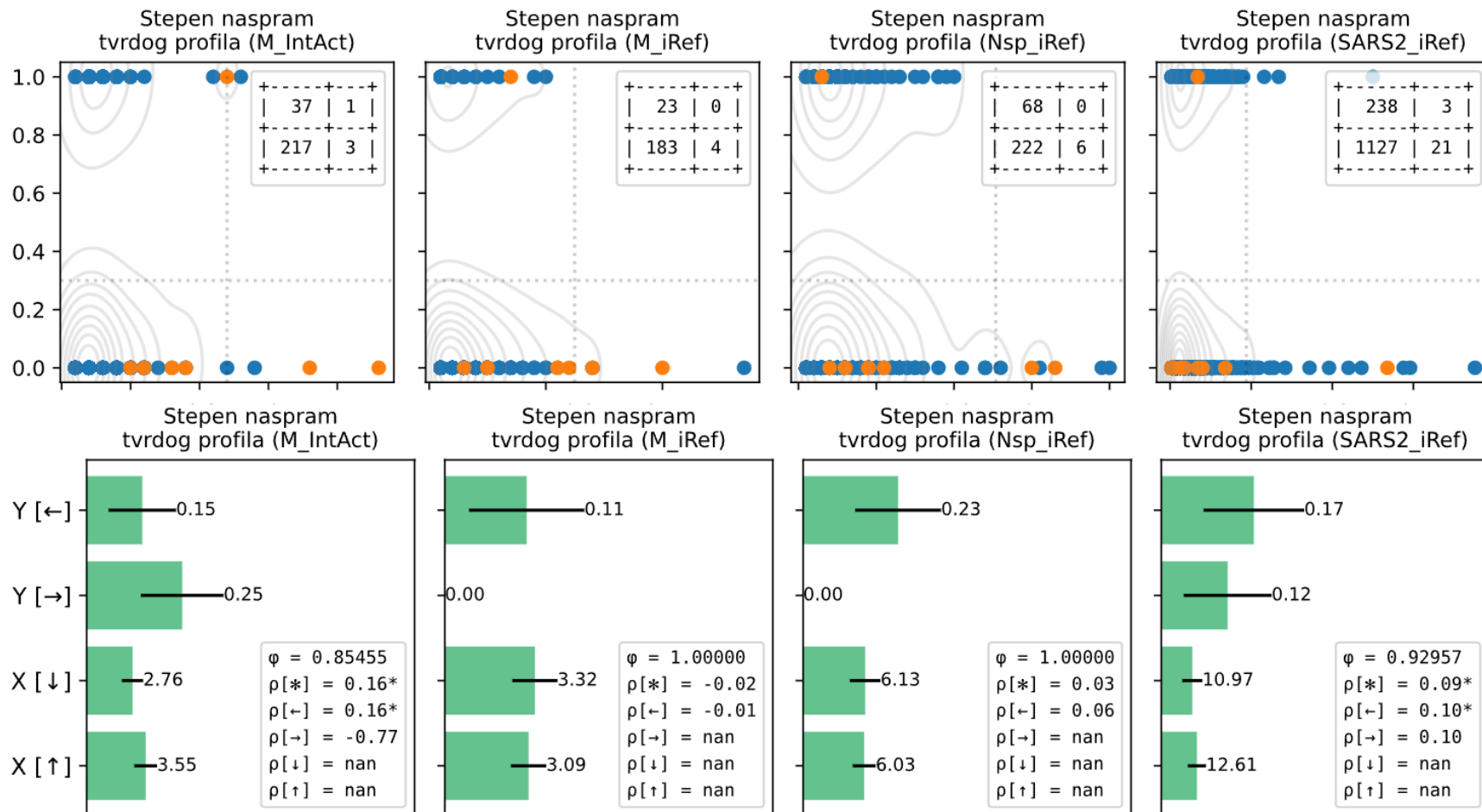
Trend raspodela sa granicama



Trend raspodela sa granicama

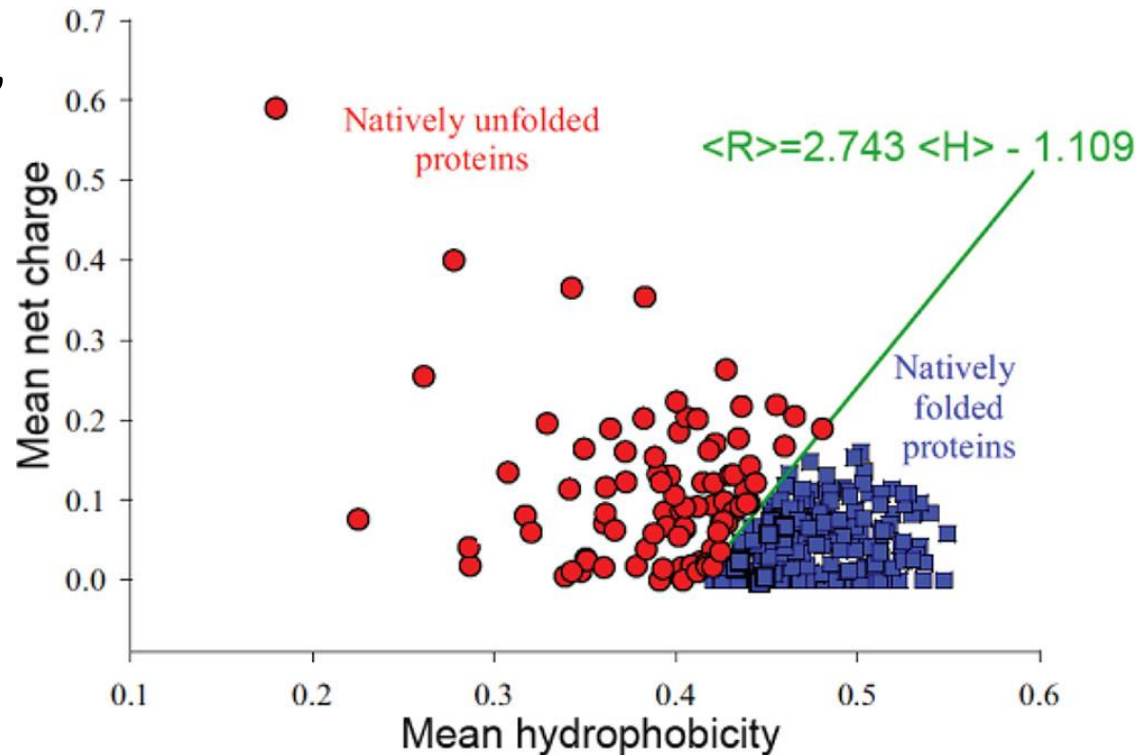


Trend raspodela sa granicama

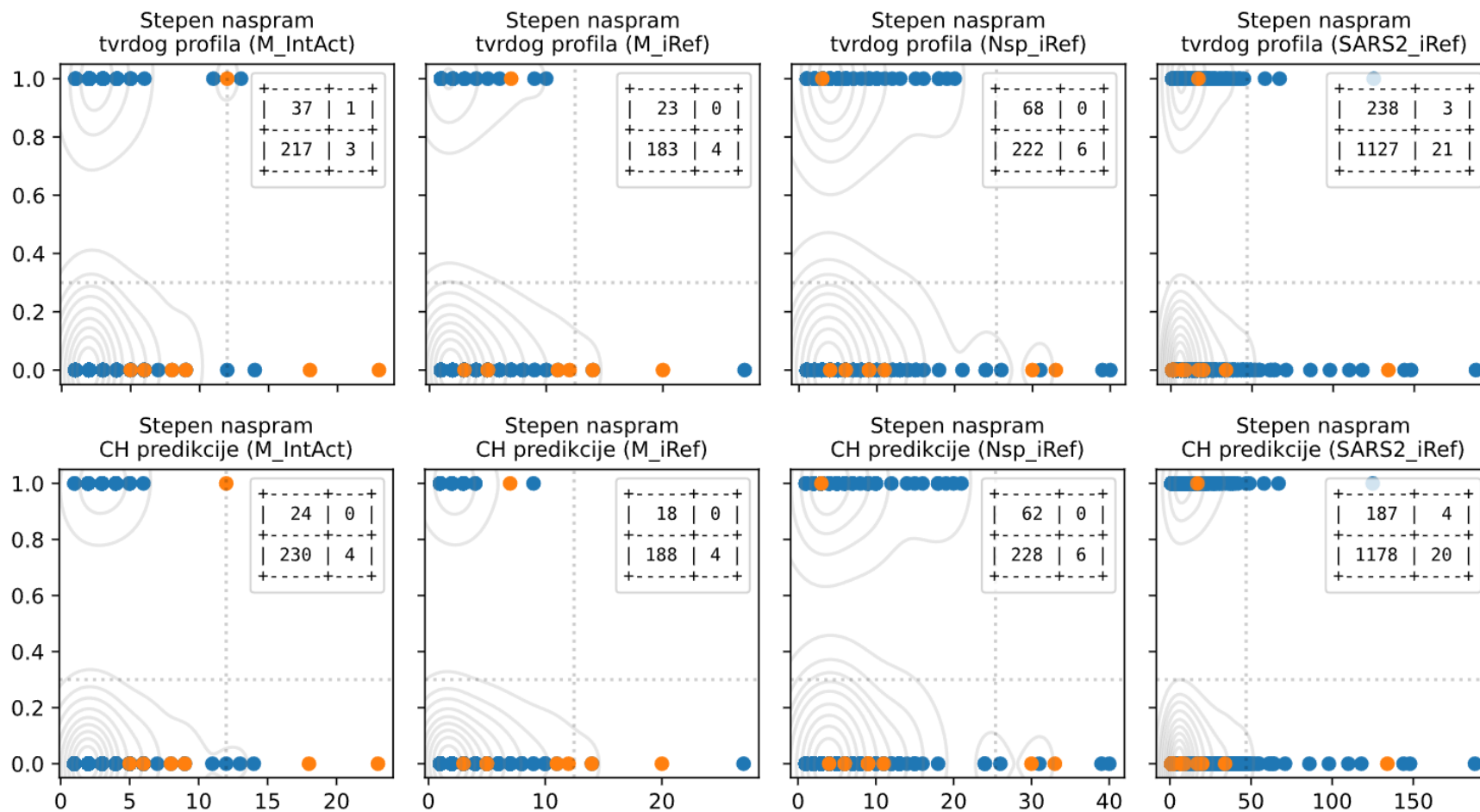


Hidropatija i *CH* predikcija

- Hidropatija \approx inverz „mekom” profilu i udelu
- Hidrofilni i naelektrisani lanci \rightarrow *IDP*
- *CH* predikcija \approx „tvrdi” profil

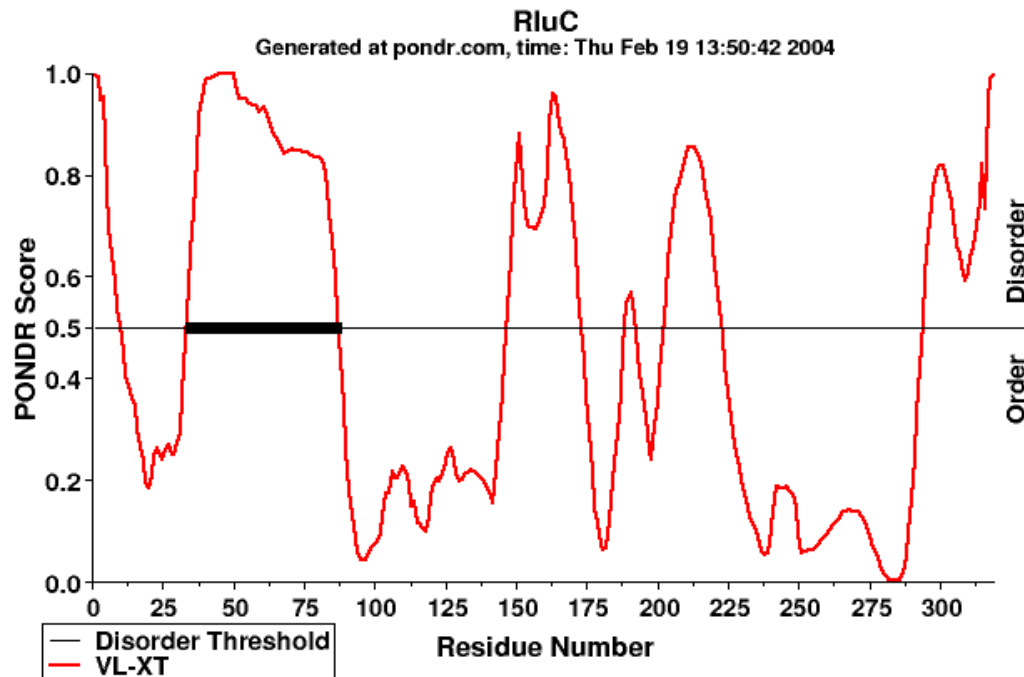


CH predikcija \approx „tvrđi” profil



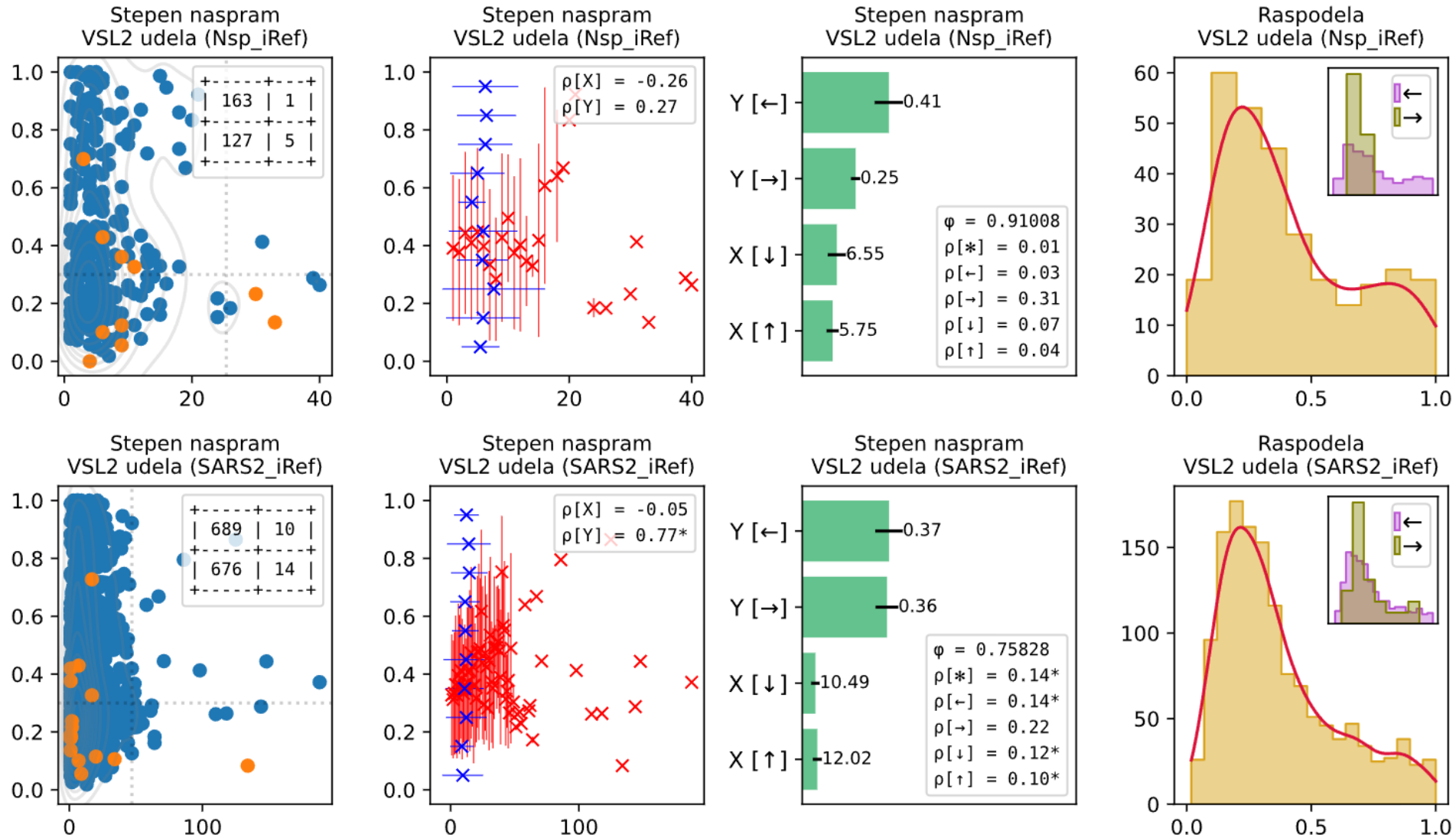
PONDR[®] familija

- Više zadataka – ansambli neuronskih mreža
- *PPIDR* – predviđanje po svakom ostatku



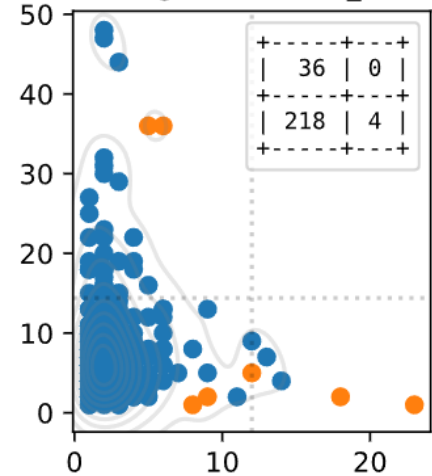
- PONDR[®], <http://www.pondr.com/pondr-tut.html>

PONDR[®] VSL2

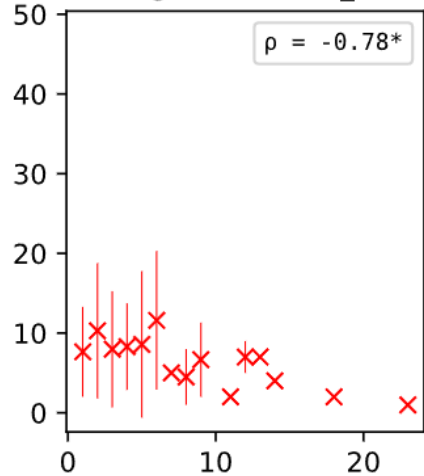


Neuređeni regioni

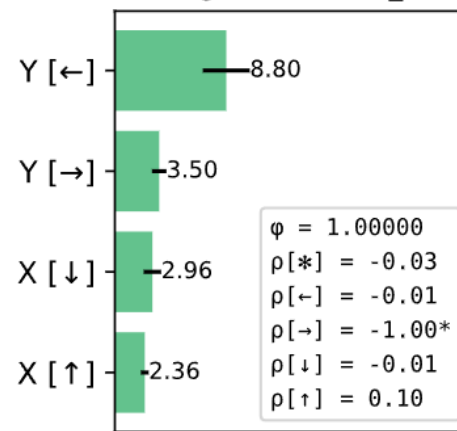
Stepen naspram
VSL2 regiona ≥ 5 (M_IntAct)



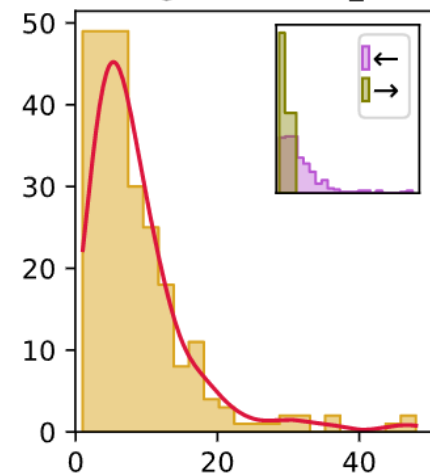
Stepen naspram
VSL2 regiona ≥ 5 (M_IntAct)



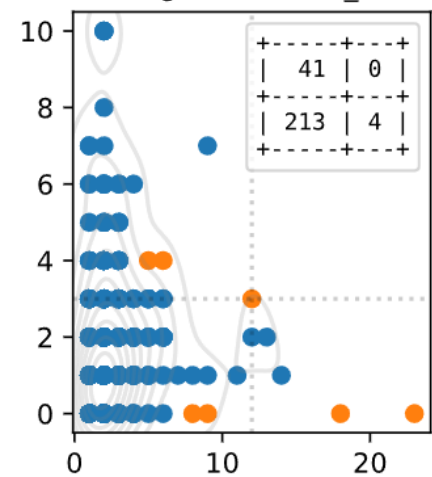
Stepen naspram
VSL2 regiona ≥ 5 (M_IntAct)



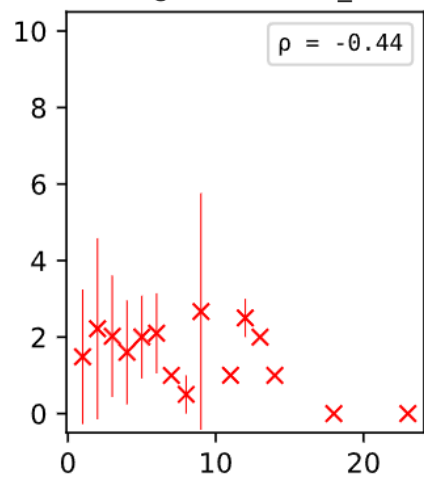
Raspodela
VSL2 regiona ≥ 5 (M_IntAct)



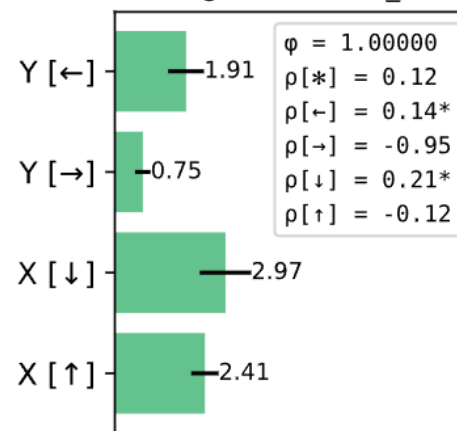
Stepen naspram
VSL2 regiona > 30 (M_IntAct)



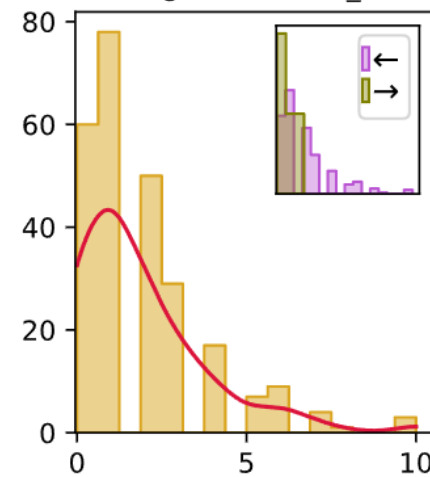
Stepen naspram
VSL2 regiona > 30 (M_IntAct)



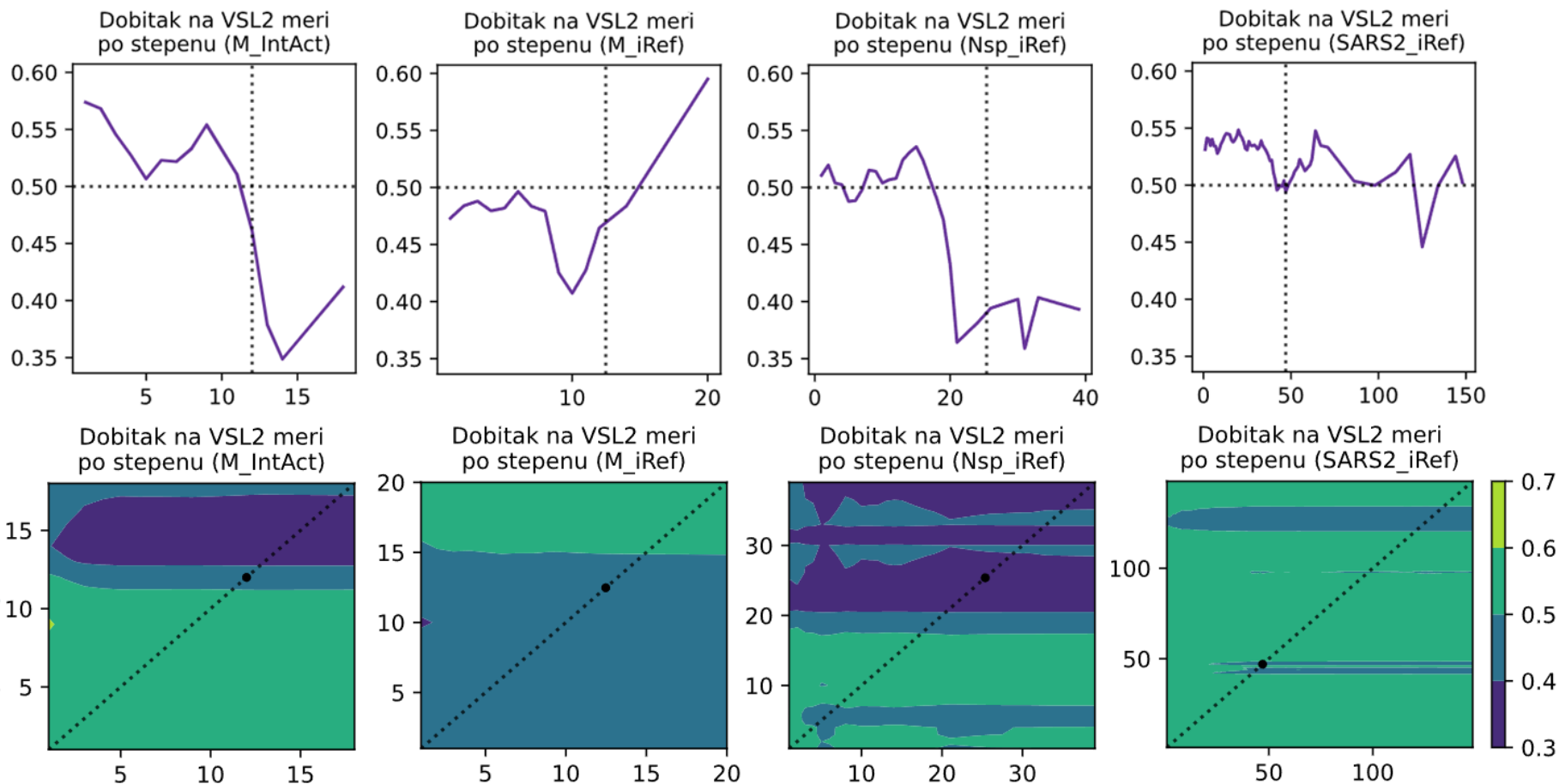
Stepen naspram
VSL2 regiona > 30 (M_IntAct)



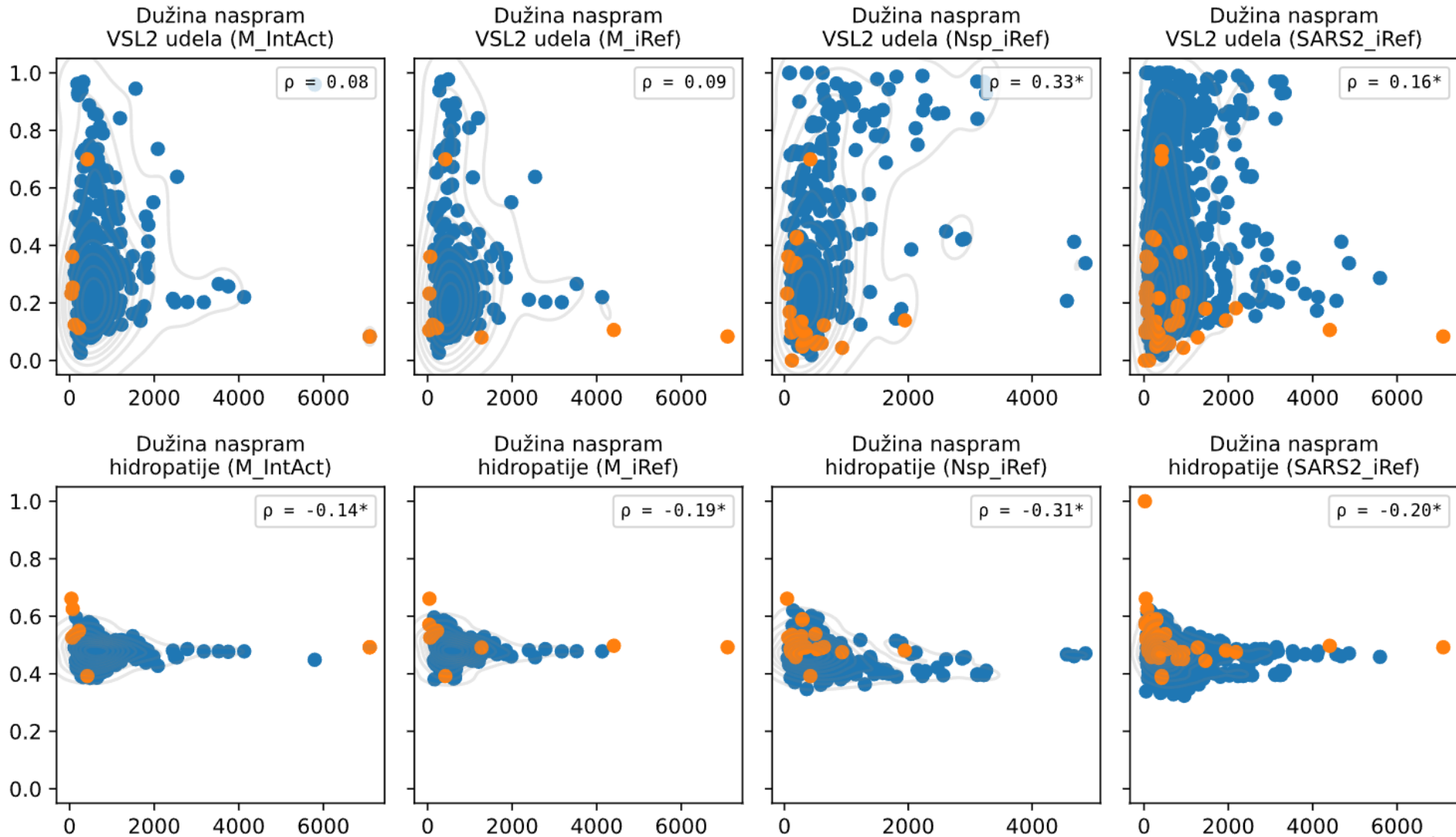
Raspodela
VSL2 regiona > 30 (M_IntAct)



Raspodela „dobitaka” prema podeli

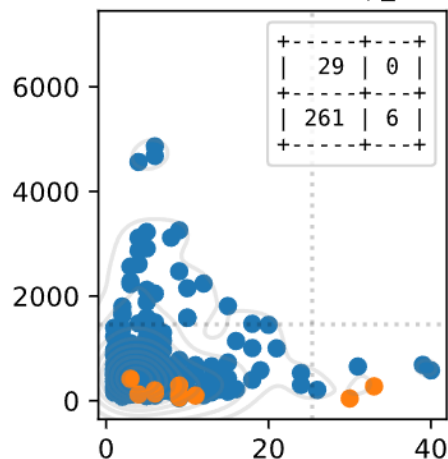


Dužina sekvence

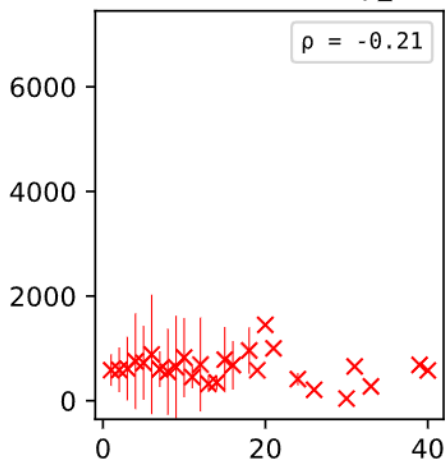


Dužina sekvence

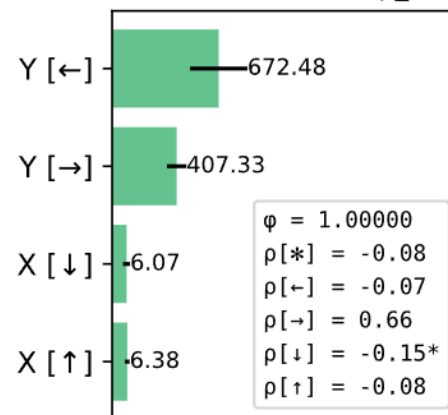
Stepen naspram
dužine sekvence (Nsp_iRef)



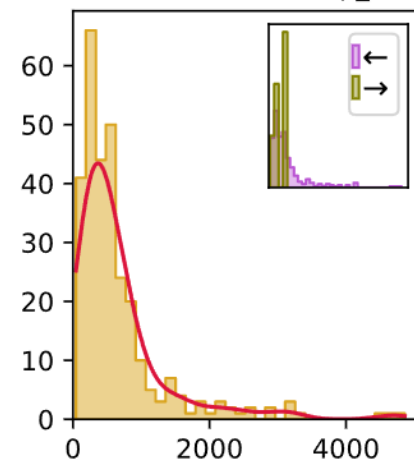
Stepen naspram
dužine sekvence (Nsp_iRef)



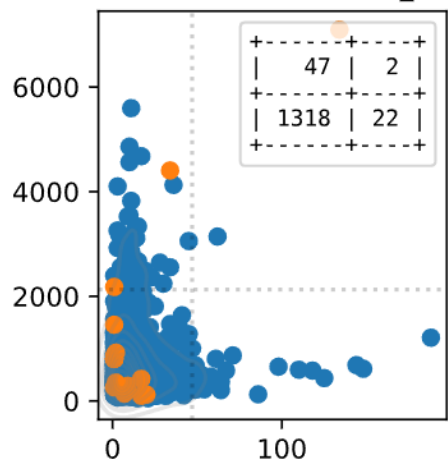
Stepen naspram
dužine sekvence (Nsp_iRef)



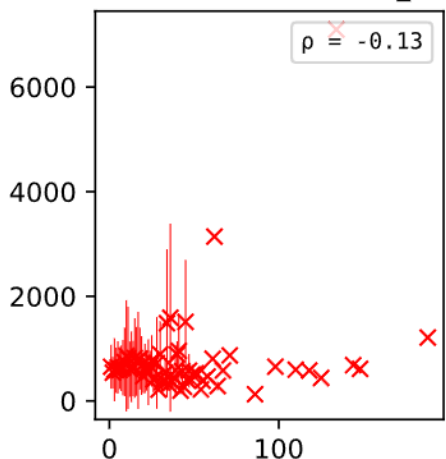
Raspodela
dužine sekvence (Nsp_iRef)



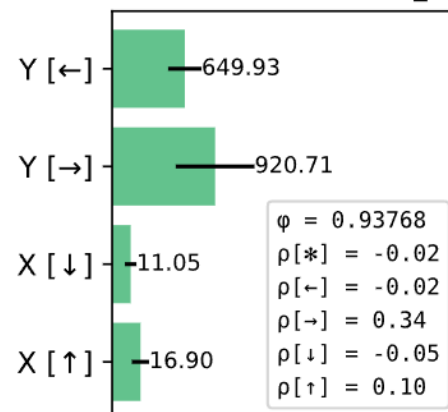
Stepen naspram
dužine sekvence (SARS2_iRef)



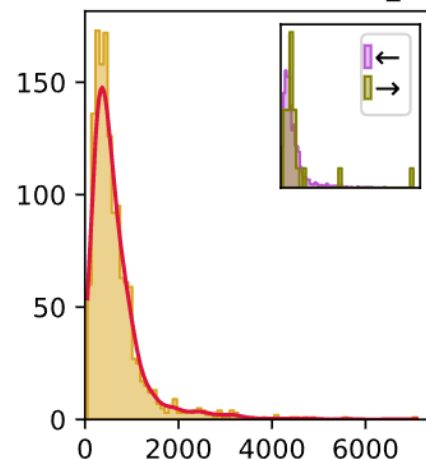
Stepen naspram
dužine sekvence (SARS2_iRef)



Stepen naspram
dužine sekvence (SARS2_iRef)

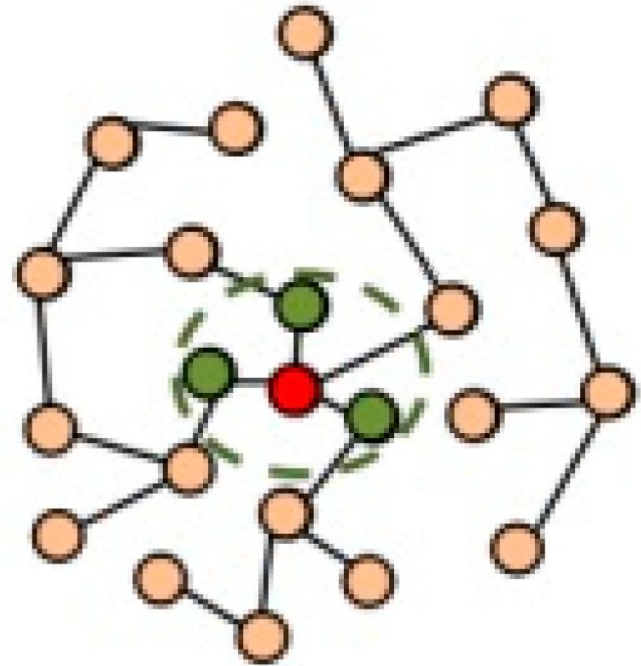


Raspodela
dužine sekvence (SARS2_iRef)

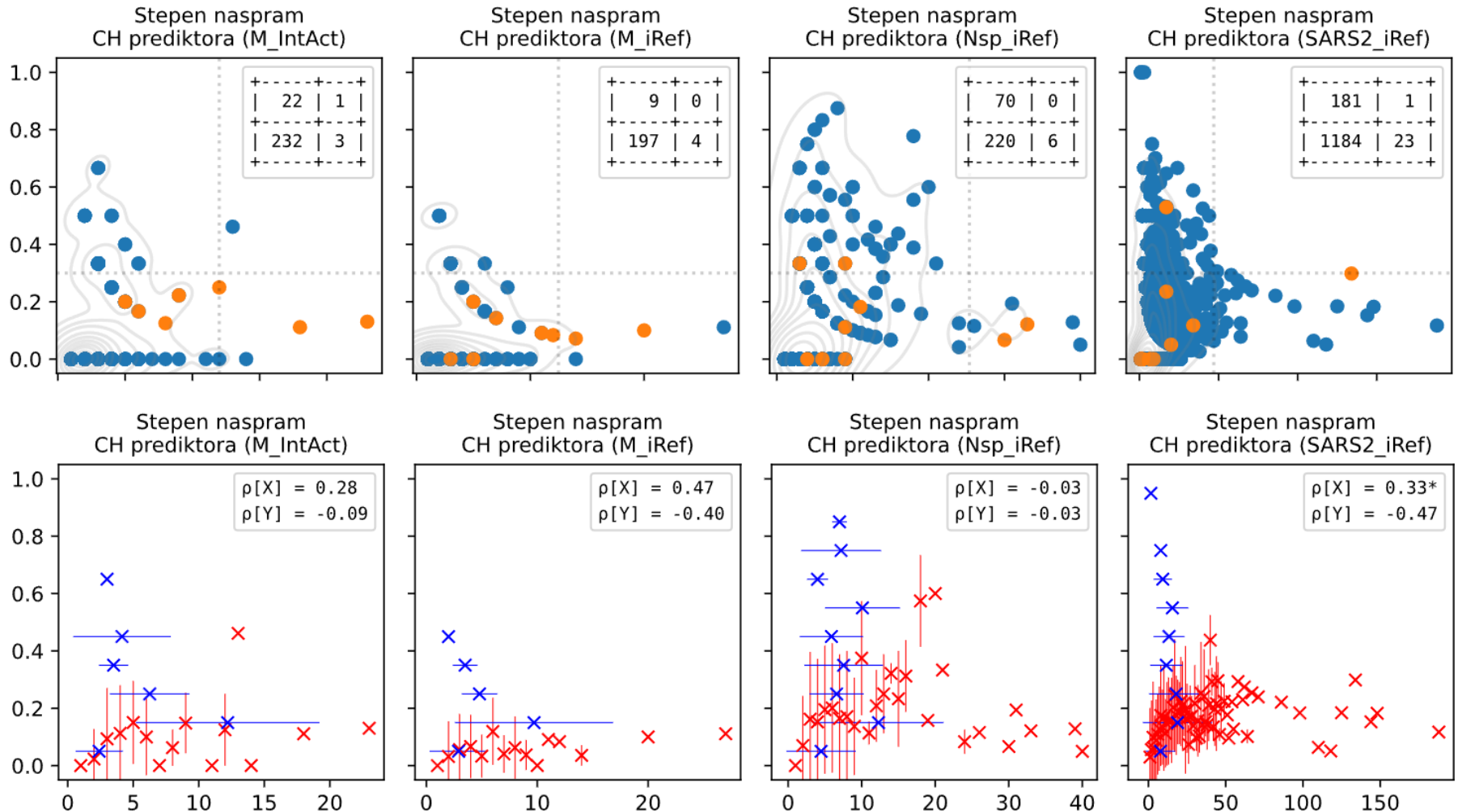


Dodatne analize

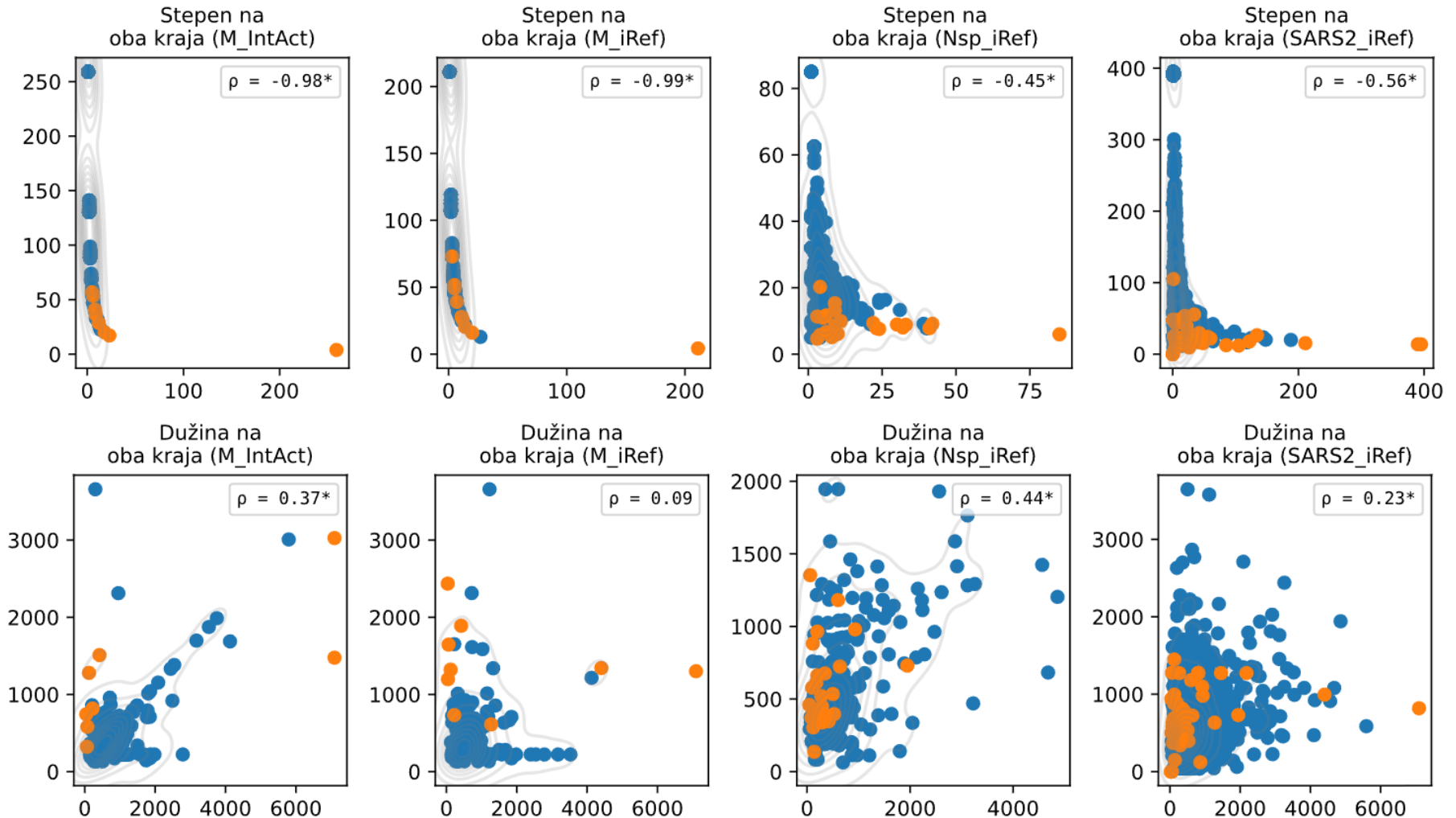
- Drugi kraj interakcije – postoji li korelacija, možda je neuređenost u susedstvu
- Analogne analize – bilo koje mere u paru
- Uzajamno dejstvo – postoji li veza sa *IDP*



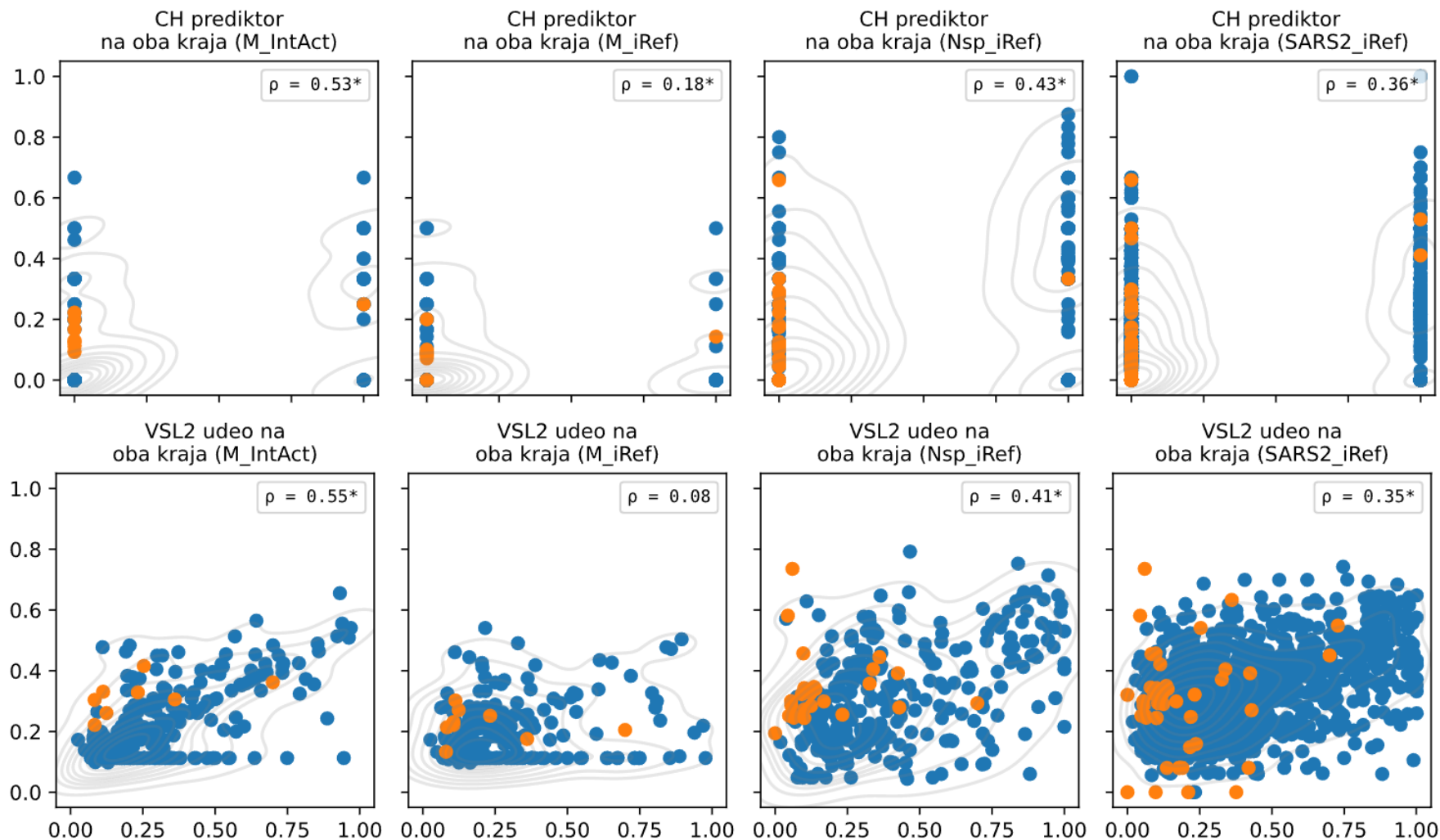
Drugi kraj interakcije



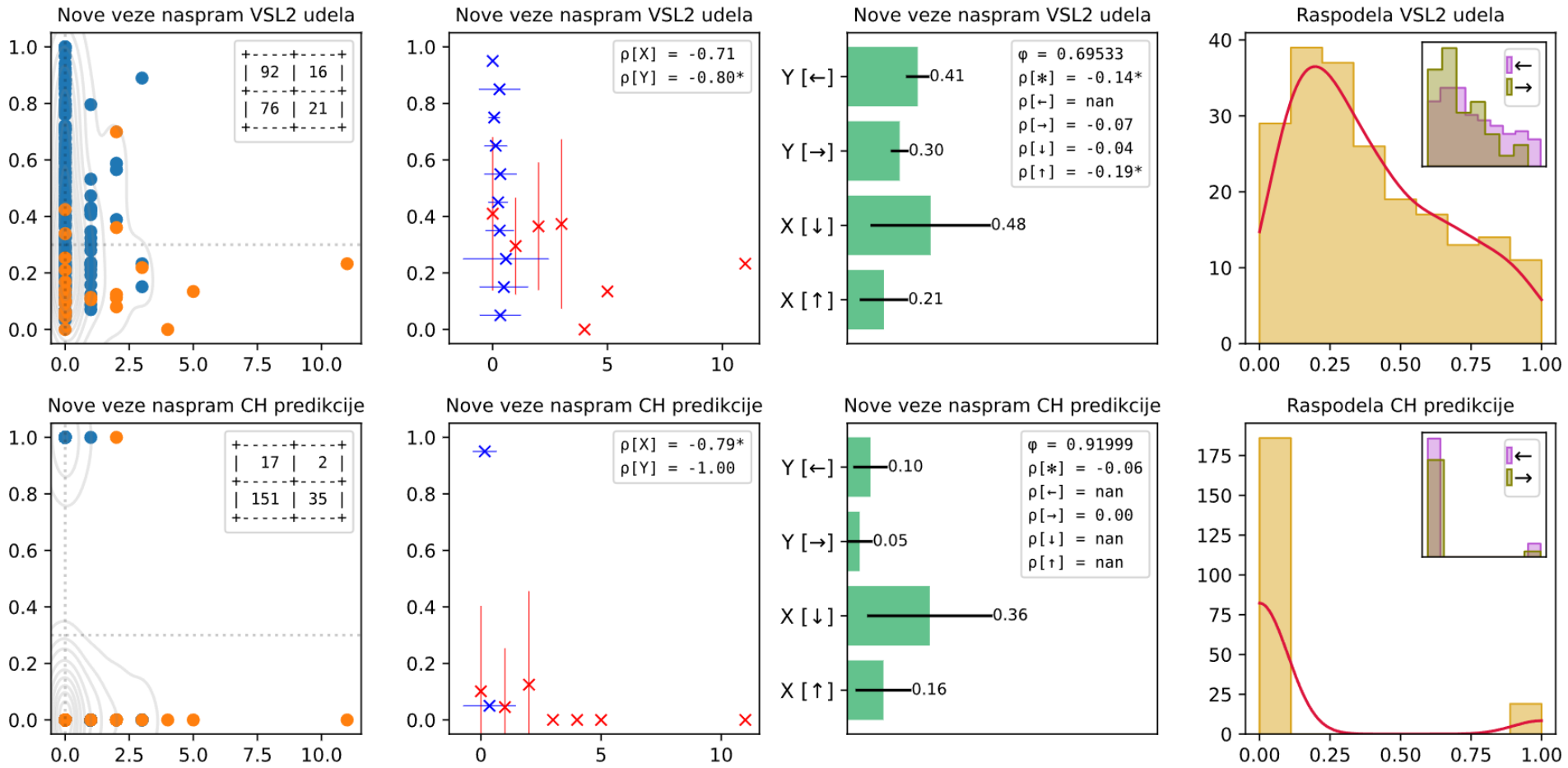
Drugi kraj interakcije



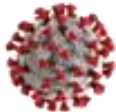

Drugi kraj interakcije



Uzajamno dejstvo



Zaključak

- Analizirano – nekoliko interaktoma virusa *SARS-CoV-2*, nekoliko mera neuređenosti, nekoliko različitih načina prikaza...
- Rezultati – iako mnogo toga zavisi od odabira mesta podele, čini se da su najpovezaniji čvorovi (čvorišta, habovi) mahom uređeniji
- Susedstvo i uzajamno dejstvo – slično
- Uzrok – virusni proteini?  

Literatura

- Vladimir N. Uversky. Analyzing IDPs in Interactomes. In Birthe B. Kragelund and Karen Skriver, editors, *Intrinsically Disordered Proteins*, chapter 46, pages 895–945. Springer Science+Business Media, LLC, 2020.
- Zsuzsanna Dosztányi, Jake Chen, A. Keith Dunker, István Simon, and Peter Tompa. Disorder and Sequence Repeats in Hub Proteins and Their Implications for Network Evolution. *Journal of Proteome Research*, 5(11):2985–2995, 2006.