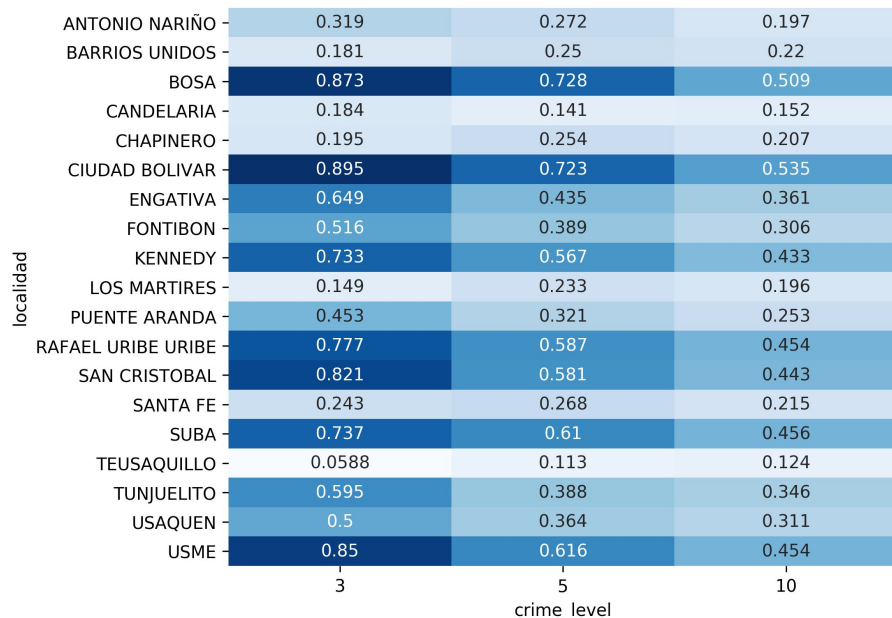# Predictability analysis RNMC dataset

March 09, 2020
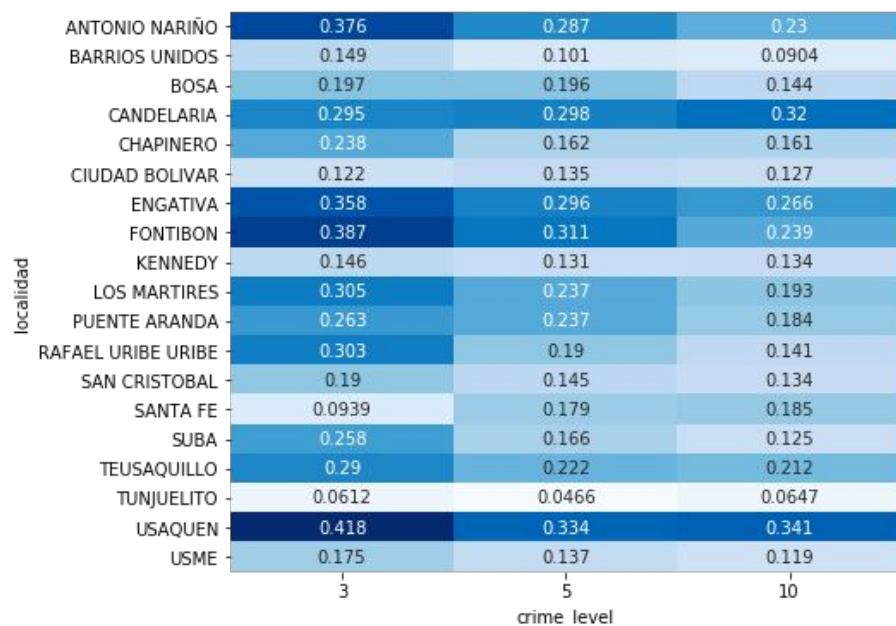
Method

- Select rnmc registers 2017-2018
- Remove outliers > q=0.99
- Assign mean values for removed points
- Compare predictability values with the ones obtained on experiment 8.

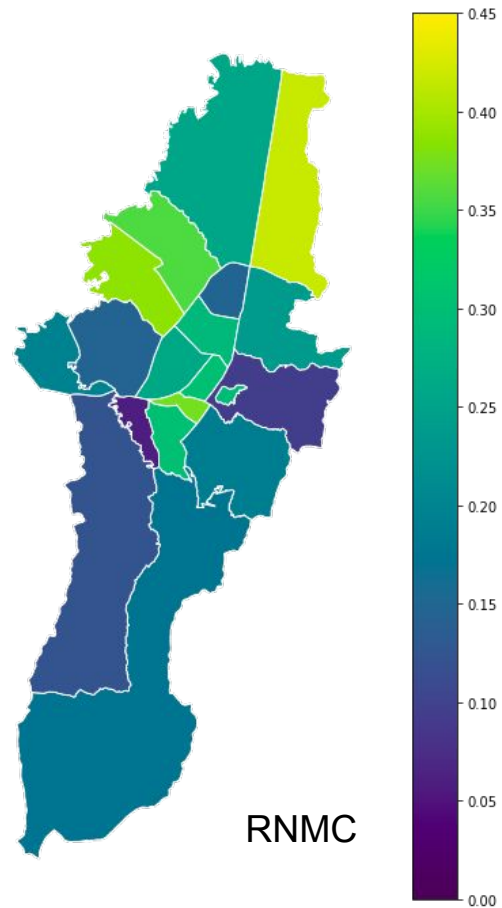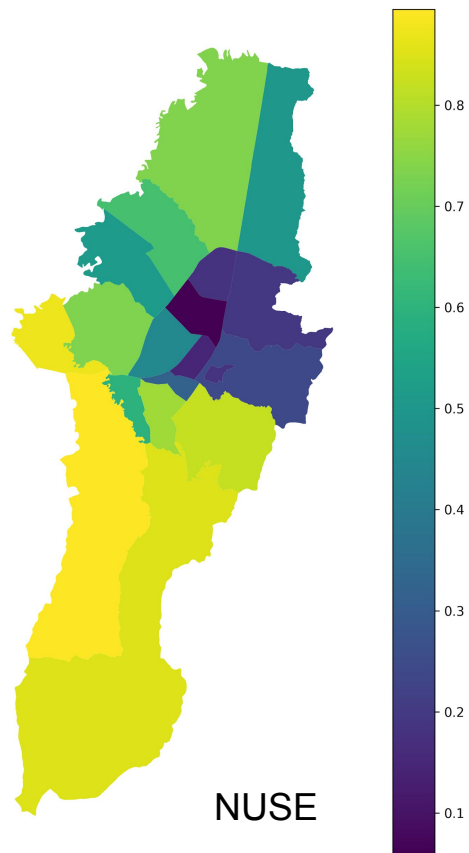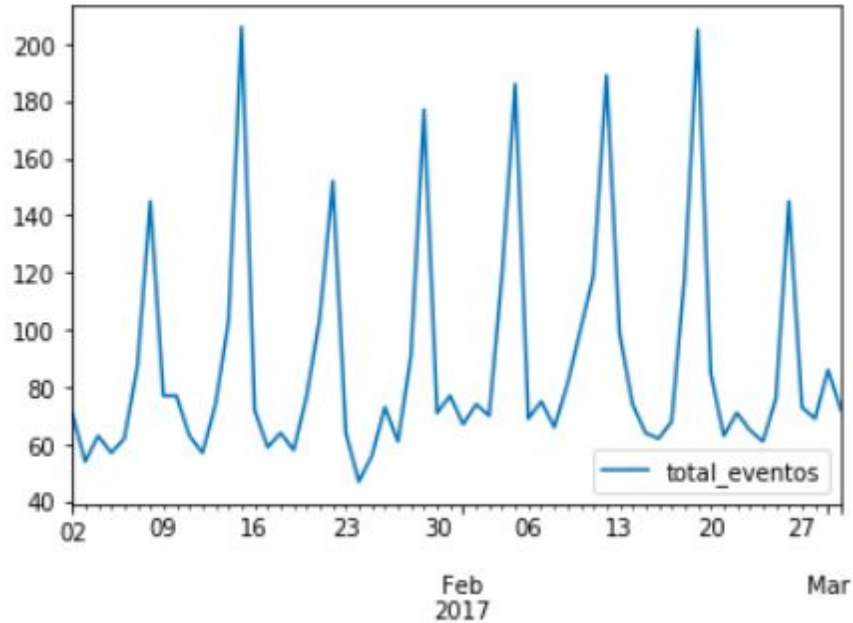# Predictability table by localidades

localidad	3	5	10
ANTONIO NARIÑO	0.319	0.272	0.197
BARRIOS UNIDOS	0.181	0.25	0.22
BOSA	0.873	0.728	0.509
CANDELARIA	0.184	0.141	0.152
CHAPINERO	0.195	0.254	0.207
CIUDAD BOLIVAR	0.895	0.723	0.535
ENGATIVA	0.649	0.435	0.361
FONTIBON	0.516	0.389	0.306
KENNEDY	0.733	0.567	0.433
LOS MARTIRES	0.149	0.233	0.196
PUENTE ARANDA	0.453	0.321	0.253
RAFAEL URIBE URIBE	0.777	0.587	0.454
SAN CRISTOBAL	0.821	0.581	0.443
SANTA FE	0.243	0.268	0.215
SUBA	0.737	0.61	0.456
TEUSAQUILLO	0.0588	0.113	0.124
TUNJUELITO	0.595	0.388	0.346
USAQUEN	0.5	0.364	0.311
USME	0.85	0.616	0.454

crime_level

NUSE

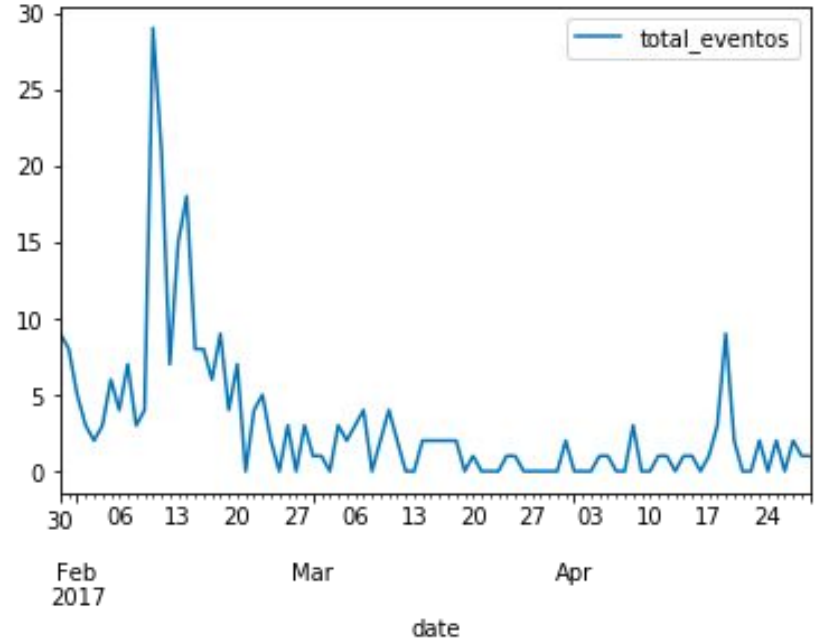| localidad	3	5	10
ANTONIO NARIÑO	0.376	0.287	0.23
BARRIOS UNIDOS	0.149	0.101	0.0904
BOSA	0.197	0.196	0.144
CANDELARIA	0.295	0.298	0.32
CHAPINERO	0.238	0.162	0.161
CIUDAD BOLIVAR	0.122	0.135	0.127
ENGATIVA	0.358	0.296	0.266
FONTIBON	0.387	0.311	0.239
KENNEDY	0.146	0.131	0.134
LOS MARTIRES	0.305	0.237	0.193
PUENTE ARANDA	0.263	0.237	0.184
RAFAEL URIBE URIBE	0.303	0.19	0.141
SAN CRISTOBAL	0.19	0.145	0.134
SANTA FE	0.0939	0.179	0.185
SUBA	0.258	0.166	0.125
TEUSAQUILLO	0.29	0.222	0.212
TUNJUELITO	0.0612	0.0466	0.0647
USAQUEN	0.418	0.334	0.341
USME	0.175	0.137	0.119

crime_level

RNMC

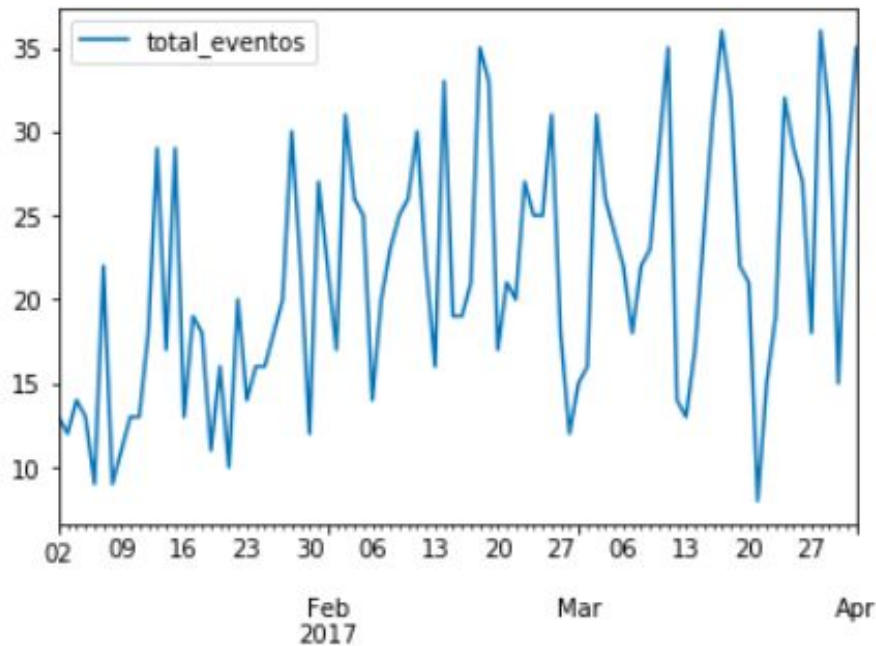Predictability map

Ciudad Bolívar time series
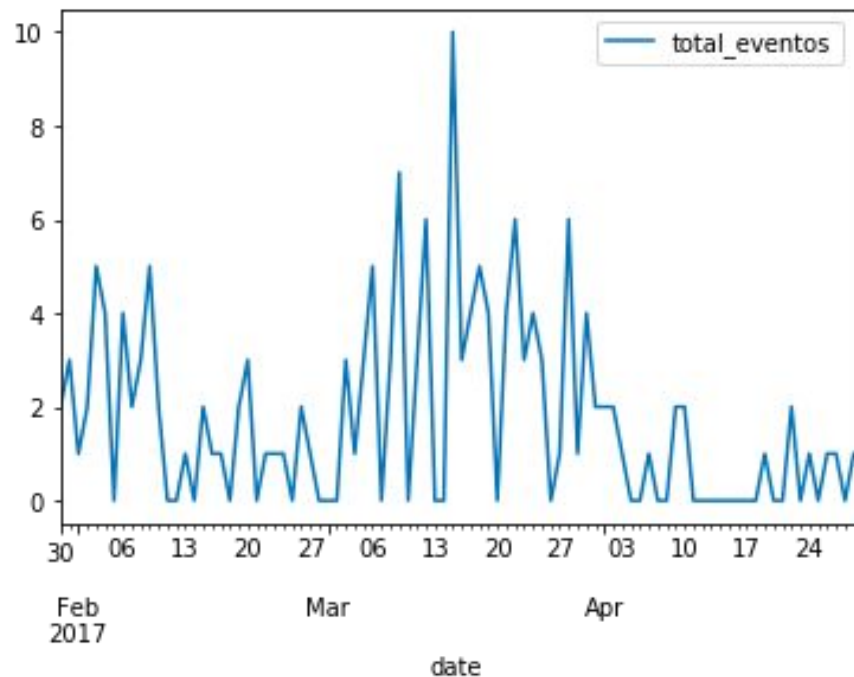

NUSE

RNMC

Teusaquillo time series

NUSE

RNMC

Conclusions

- Predictability analysis by localidades using RNMC dataset doesn't match with the reported using NUSE dataset.
- Aggressive behavior penalty fees might be affected by other factors related to the penalty application dynamics (e.g. resources available to penalties application)
- Data quality regarding Colwell's methodology and time series analysis. High proportion of days without data (non-consecutive dates)