

In [2]:

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Wed Dec  2 19:06:34 2020

@author: nacho

crear para todos los estados dataset con diccionario

"""
#%%
import statsmodels.api as sm
import pandas as pd
import numpy as np
import matplotlib
import seaborn as sns; sns.set()
from scipy import stats
from matplotlib import pyplot as plt
sns.set(color_codes=True)
import os, datetime
from collections import Counter
from matplotlib.offsetbox import AnchoredText
import matplotlib.ticker as ticker
from darts import TimeSeries
from darts.metrics import mape, mase
from darts.utils.statistics import check_seasonality, plot_acf, plot_residuals_a
from darts.models import (
    NaiveSeasonal,
    NaiveDrift,
    Prophet,
    ExponentialSmoothing,
    ARIMA,
    AutoARIMA,
    StandardRegressionModel,
    Theta,
    FFT
)
import warnings
warnings.filterwarnings("ignore")
import logging
logging.disable(logging.CRITICAL)
import statsmodels.api as sm
from itertools import product
import math
from sklearn.preprocessing import MinMaxScaler
matplotlib.rcParams['figure.figsize'] = [9.0, 5.0]
```

In [4]:

```
path = "/home/nacho/Documents/coronavirus/Data_Mexico"
os.chdir(os.path.join(path))
delitos_df = pd.read_csv("data/IDEFC_NM_dic2020.csv", encoding='latin-1',thousar
#delitos_df = pd.read_csv("data/IDEFC_NM_mar2021.csv", encoding='latin-1',thousa
#movilidad = movilidad.return_df()
df_movilidad = pd.read_csv("data/google_movilidad.csv", index_col='Date')
```

In [9]:

```
delitos_df.head()
```

Out[9]:

	Año	Clave_Ent	Entidad	Bien jurídico afectado	Tipo de delito	Subtipo de delito	Modalidad	Enero	Febrero	Marz
0	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	Con arma de fuego	3	0	
1	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	Con arma blanca	1	1	
2	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	Con otro elemento	0	0	
3	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio doloso	No especificado	2	0	
4	2015	1	Aguascalientes	La vida y la Integridad corporal	Homicidio	Homicidio culposo	Con arma de fuego	0	0	

In [10]:

```
df_movilidad.head()
```

Out[10]:

	Unnamed: 0	State	Residences
Date			
2020-02-16	2667462	NaN	-1.0
2020-02-17	2667463	NaN	-2.0
2020-02-18	2667464	NaN	-2.0
2020-02-19	2667465	NaN	-1.0
2020-02-20	2667466	NaN	-1.0

In [5]:

```
#####
#####ESTADOS#####
#####
```

In [6]:

```
##%hacer listas de valores unicos
dict_delitos_sum = {}
states = delitos_df.Entidad.value_counts().index.tolist()
delitos_list = delitos_df['Tipo de delito'].value_counts().index.tolist()
```

In [11]:

```
delitos_list
```

Out[11]:

```
['Robo',
 'Homicidio',
 'Lesiones',
```

```

'Secuestro',
'Feminicidio',
'Extorsión',
'Otros delitos que atentan contra la libertad y la seguridad sexual',
'Otros delitos contra la familia',
'Rapto',
'Falsedad',
'Incumplimiento de obligaciones de asistencia familiar',
'Incesto',
'Violencia familiar',
'Otros delitos que atentan contra la vida y la integridad corporal',
'Despojo',
'Aborto',
'Delitos cometidos por servidores públicos',
'Tráfico de menores',
'Otros delitos del Fuero Común',
'Abuso de confianza',
'Otros delitos que atentan contra la libertad personal',
'Contra el medio ambiente',
'Trata de personas',
'Violación simple',
'Narcomenudeo',
'Otros delitos contra la sociedad',
'Otros delitos contra el patrimonio',
'Falsificación',
'Violación equiparada',
'Acoso sexual',
'Fraude',
'Daño a la propiedad',
'Evasión de presos',
'Hostigamiento sexual',
'Corrupción de menores',
'Abuso sexual',
'Amenazas',
'ElectORAles',
'Allanamiento de morada',
'Violencia de género en todas sus modalidades distinta a la violencia familia
r']

```

```

In [12]: ###Generar dict de delitos por estado
for state in states:
    delitos_state = delitos_df[delitos_df['Entidad'] == state]
    delitos_sum_df = pd.DataFrame(index = pd.date_range('2015-01', '2021-01', fr
    #delitos_sum_df = pd.DataFrame(index = pd.date_range('2015-01', '2021-04', i
    delitos_sum_df.index.name = 'date'
    years = np.arange(2015,2021)
    #years = np.arange(2015,2022)
    meses = ['Enero', 'Febrero', 'Marzo', 'Abril', 'Mayo', 'Junio', 'Julio', 'Ag
    for delito in delitos_list:
        valores_delito = []
        for year in years:
            for mes in meses:
                value = delitos_state.loc[(delitos_state['Tipo de delito'] == de
                valores_delito.append(value)
            delitos_sum_df[str(delito)] = valores_delito
        dict_delitos_sum[state] = delitos_sum_df

```

```

In [16]: dict_delitos_sum['Colima']['Robo']

```

```

Out[16]: date
2015-01-31    188

```

```

2015-02-28    209
2015-03-31    213
2015-04-30    200
2015-05-31    213
...
2020-08-31    500
2020-09-30    639
2020-10-31    705
2020-11-30    560
2020-12-31    566
Freq: M, Name: Robo, Length: 72, dtype: int64

```

```

In [17]: ##generar dict con todas las comb de estados y delitos para guardar corr y trend
df_dsct = pd.DataFrame(list(product(states, delitos_list)), columns=['state', 'delito', 'corr', 'trend'])
df_dsct['corr'] = ""
df_dsct['trend'] = ""

```

```

In [18]: df_dsct.head()

```

```

Out[18]:
   state  delito  corr  trend
0  Quintana Roo    Robo
1  Quintana Roo  Homicidio
2  Quintana Roo  Lesiones
3  Quintana Roo  Secuestro
4  Quintana Roo  Femicidio

```

```

In [19]: ##Generar df por delito por estado y su porcentaje de crecimiento respecto al año anterior
for state in states:
    for delito in delitos_list:
        df_delito_state = dict_delitos_sum[state].sort_values(['date'])[delito]
        trend = df_delito_state.resample('Y').sum().pct_change()[-1]
        if np.isinf(trend):
            trend = math.inf
        df_dsct.loc[(df_dsct['state'] == state) & (df_dsct['delito'] == delito), 'trend'] = trend

```

```

In [20]: df_dsct.head()

```

```

Out[20]:
   state  delito  corr  trend
0  Quintana Roo    Robo    -0.226434
1  Quintana Roo  Homicidio    -0.0906667
2  Quintana Roo  Lesiones    -0.0885767
3  Quintana Roo  Secuestro    -0.541667
4  Quintana Roo  Femicidio    -0.0625

```

```

In [21]: ##corr con moviliad
df_corr_movilidad = df_movilidad[df_movilidad['State'].notna()]
df_corr_movilidad['State'].replace({
    'Nuevo Leon': 'Nuevo León',
})

```

```

'Mexico City': 'Ciudad de México',
'Yucatan': 'Yucatán',
'Veracruz': 'Veracruz de Ignacio de la Llave',
'San Luis Potosí': 'San Luis Potosí',
'workplaces_percent_change_from_baseline': 'Querétaro',
'Michoacán': 'Michoacán de Ocampo',
'State of Mexico': 'México',
'Coahuila': 'Coahuila de Zaragoza'
}, inplace=True)

for state in states:
    for delito in delitos_list:
        df_delito_state = dict_delitos_sum[state].sort_values(['date'])[delito]
        movilidad_spec = df_corr_movilidad.loc[(df_corr_movilidad['State'] == state) & (df_corr_movilidad['date'] == df_delito_state['date'])]
        movilidad_spec.index = pd.to_datetime(movilidad_spec.index)
        movilidad_spec = movilidad_spec['Residences'].resample('M').sum()
        df_delito_state = df_delito_state.loc['2020-02-01':]
        corr = stats.pearsonr(movilidad_spec, df_delito_state)[0]
        df_dsct.loc[(df_dsct['state'] == state) & (df_dsct['delito'] == delito), 'corr'] = corr
        df_dsct.loc[(df_dsct['state'] == state) & (df_dsct['delito'] == delito), 'trend'] = df_delito_state['date'].max()

```

In [22]: `df_dsct.head()`

Out[22]:

	state	delito	corr	trend
0	Quintana Roo	Robo	-0.894344	-0.226434
1	Quintana Roo	Homicidio	-0.515873	-0.0906667
2	Quintana Roo	Lesiones	-0.662067	-0.0885767
3	Quintana Roo	Secuestro	-0.129394	-0.541667
4	Quintana Roo	Feminicidio	-0.353383	-0.0625

In [23]:

```

#%%Limpiar df
df_dsct.drop(df_dsct[df_dsct['trend'] == "inf"].index, inplace = True)
df_dsct.drop(df_dsct[df_dsct['trend'] == "nan"].index, inplace = True)
df_dsct.drop(df_dsct[df_dsct['corr'] == "inf"].index, inplace = True)
df_dsct.drop(df_dsct[df_dsct['corr'] == "nan"].index, inplace = True)
df_dsct = df_dsct[pd.notnull(df_dsct['trend'])]
df_dsct = df_dsct[pd.notnull(df_dsct['corr'])]
#drop index
df_dsct = df_dsct.sort_values(by=['corr'], ignore_index=True, ascending = False)

```

In [27]: `df_dsct`

Out[27]:

	state	delito	corr	trend
0	Baja California	Trata de personas	0.722332	1.84211
1	México	Violencia de género en todas sus modalidades d...	0.67381	0.387029
2	Jalisco	Aborto	0.671945	1.6
3	México	Violencia familiar	0.64436	0.828846
4	México	Otros delitos que atentan contra la vida y la ...	0.637117	0.270175
...	...	...	...	...

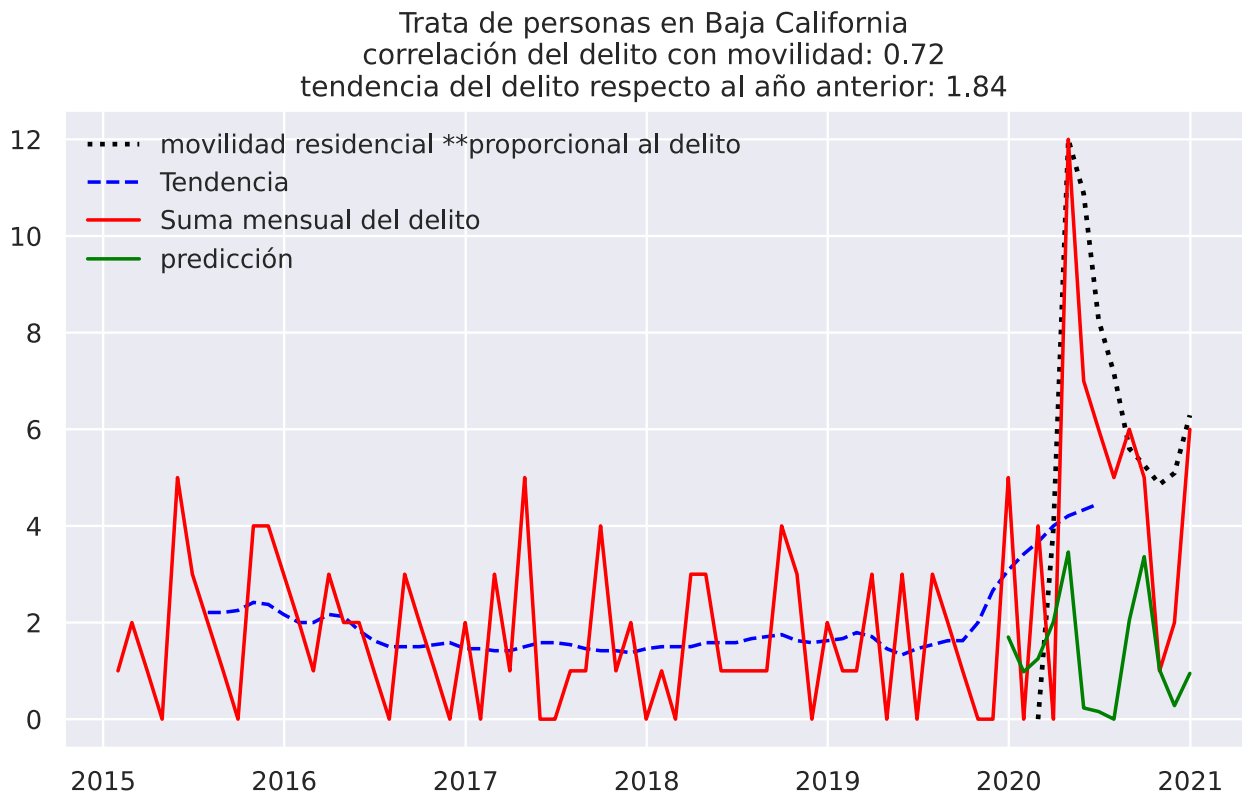


```

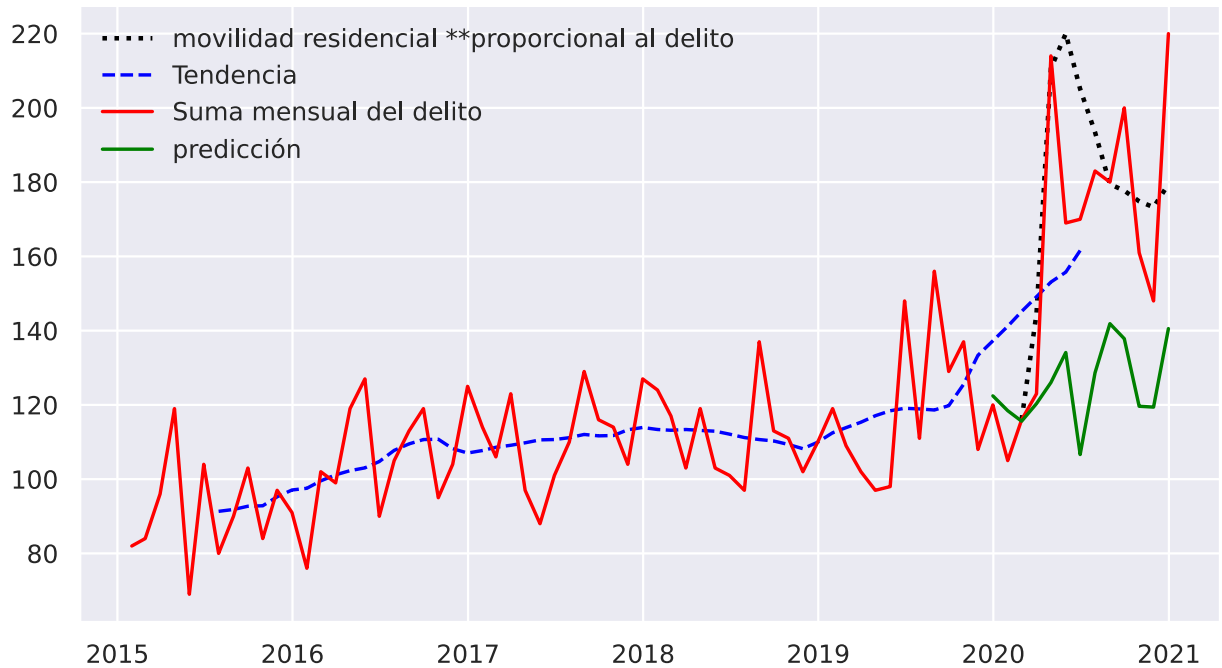
        color = 'black',
        linewidth=2)
ax.plot(decom.trend.index, decom.trend.values, label = 'Tendencia', linestyle='dashed',
ax.plot(decom.trend.index, df_delito_state.resample('M').sum(), label = 'Suma mensual del delito',
ax.set_title(delito+' en '+state+
              '\n'+correlación del delito con movilidad: '+str(round(df_dscorrel, 2))
              '\n'+tendencia del delito respecto al año anterior: '+str(round(df_dscorrel, 2))
df_pred = predict_series(df_delito_state, delito)
ax.plot(df_pred.index,
        df_pred.values,
        label = 'predicción',
        color = 'green')

ax.legend(loc='upper left')
plt.plot()

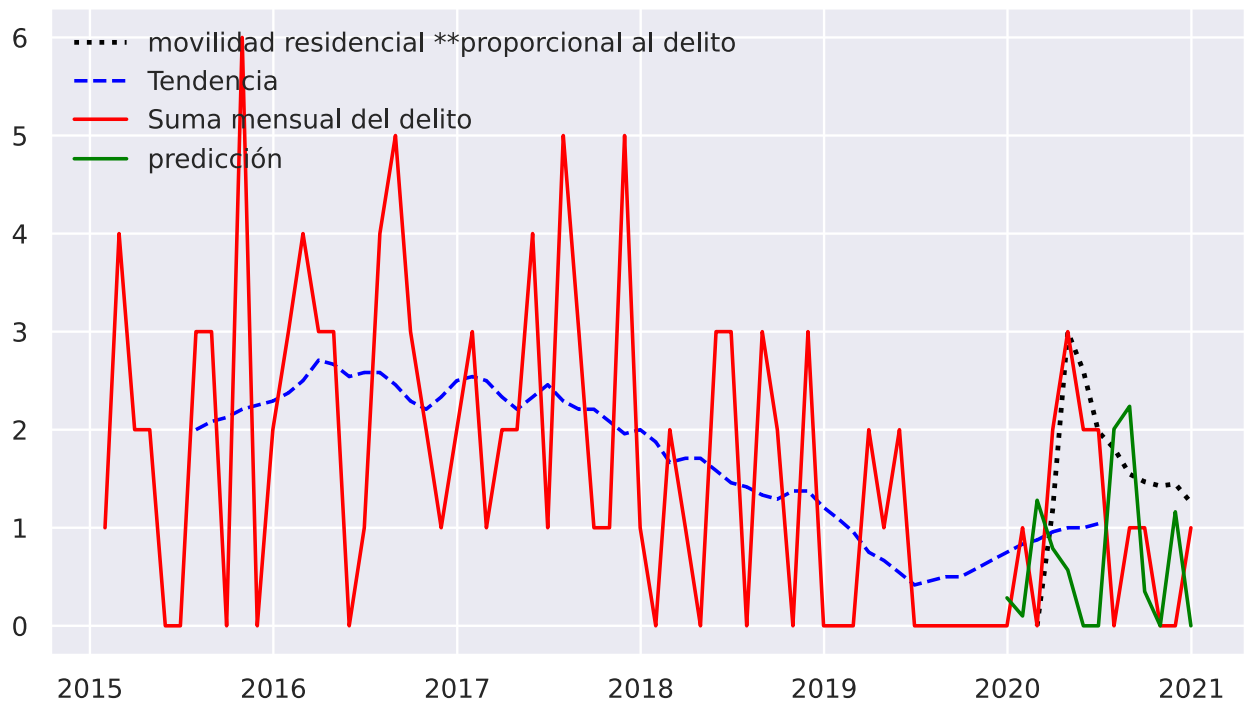
```



Violencia de género en todas sus modalidades distinta a la violencia familiar en México  
 correlación del delito con movilidad: 0.67  
 tendencia del delito respecto al año anterior: 0.39

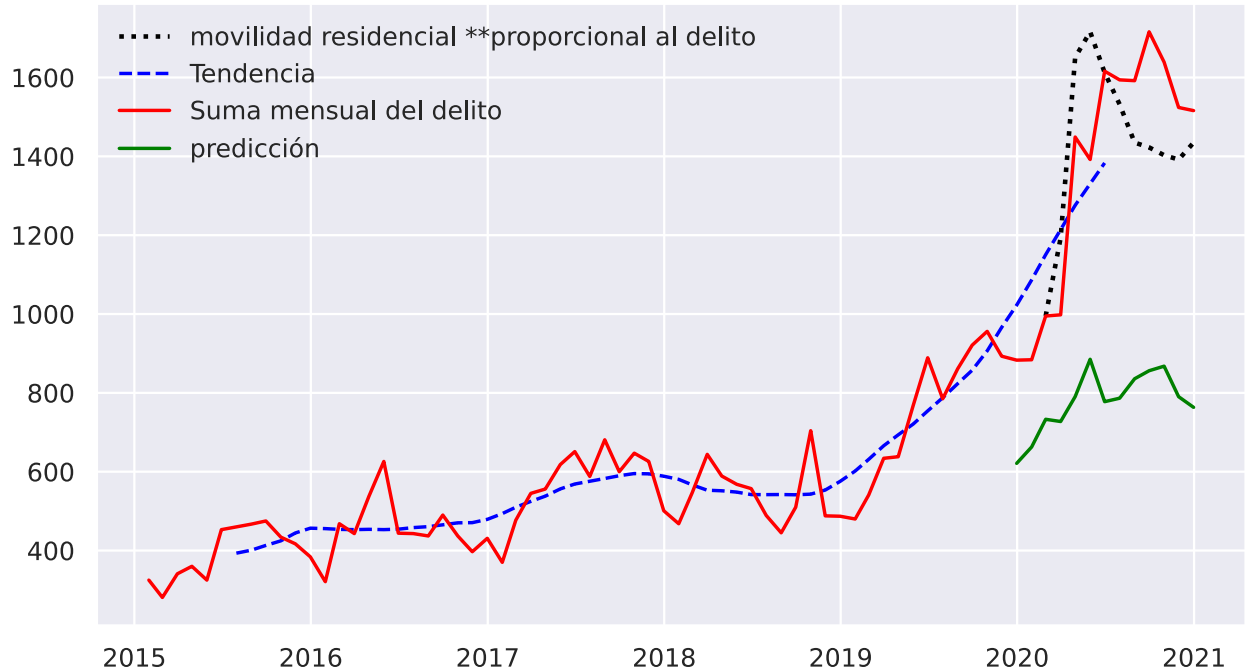


Aborto en Jalisco  
 correlación del delito con movilidad: 0.67  
 tendencia del delito respecto al año anterior: 1.6

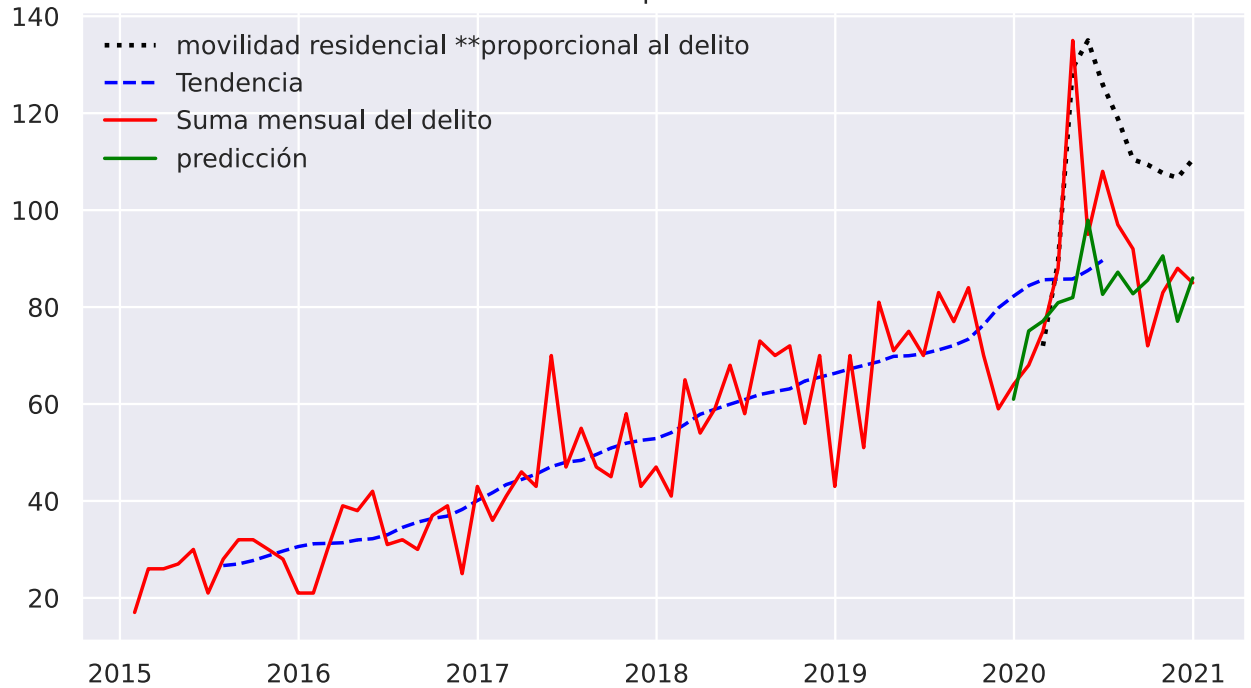




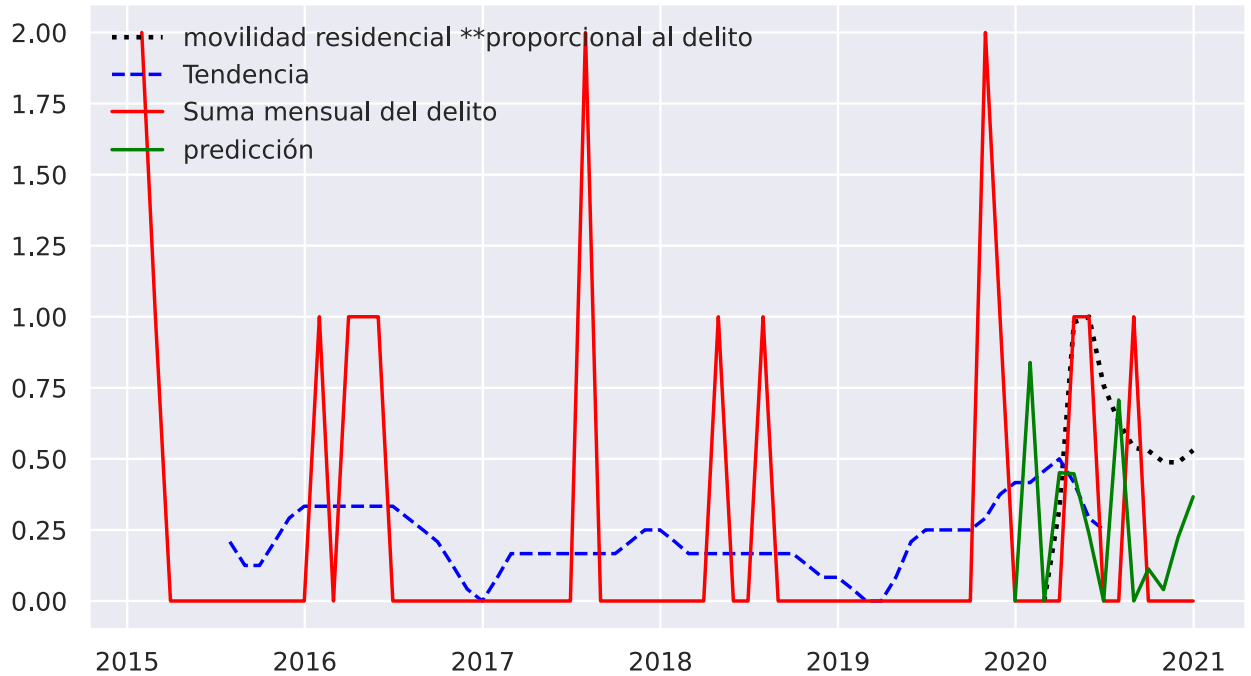
Violencia familiar en México  
 correlación del delito con movilidad: 0.64  
 tendencia del delito respecto al año anterior: 0.83



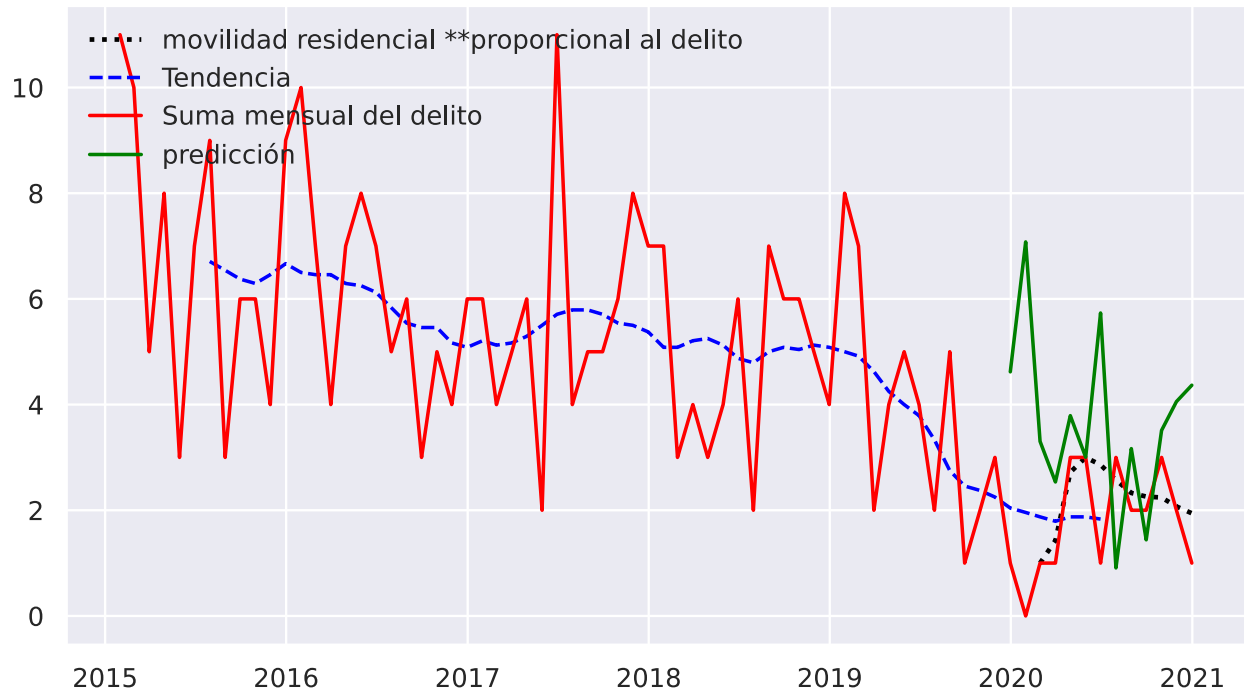
Otros delitos que atentan contra la vida y la integridad corporal en México  
 correlación del delito con movilidad: 0.64  
 tendencia del delito respecto al año anterior: 0.27



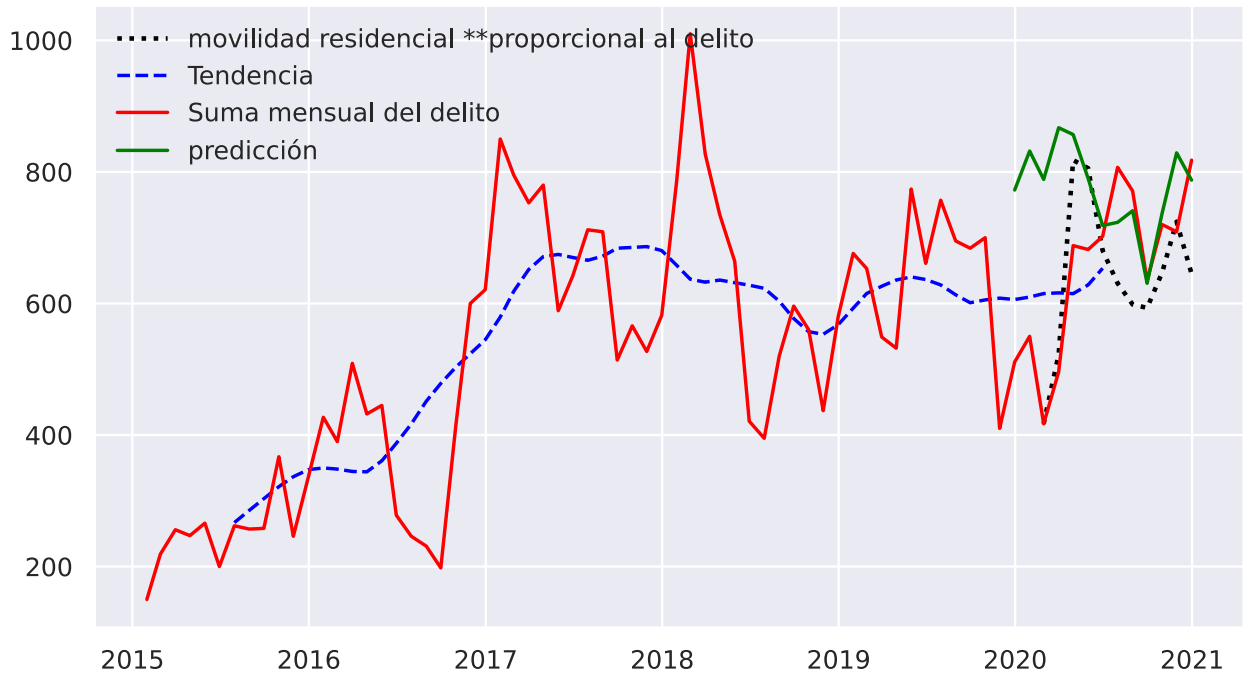
Contra el medio ambiente en Querétaro  
correlación del delito con movilidad: 0.62  
tendencia del delito respecto al año anterior: 0.0



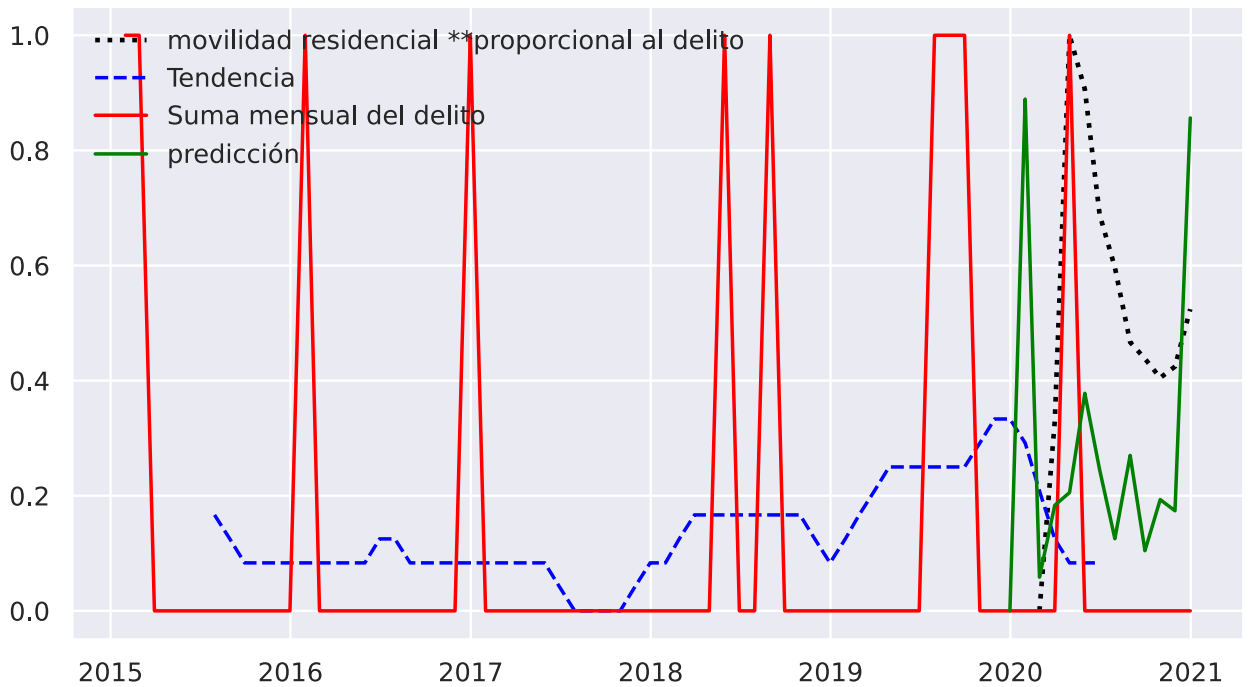
Secuestro en Guerrero  
correlación del delito con movilidad: 0.61  
tendencia del delito respecto al año anterior: -0.5



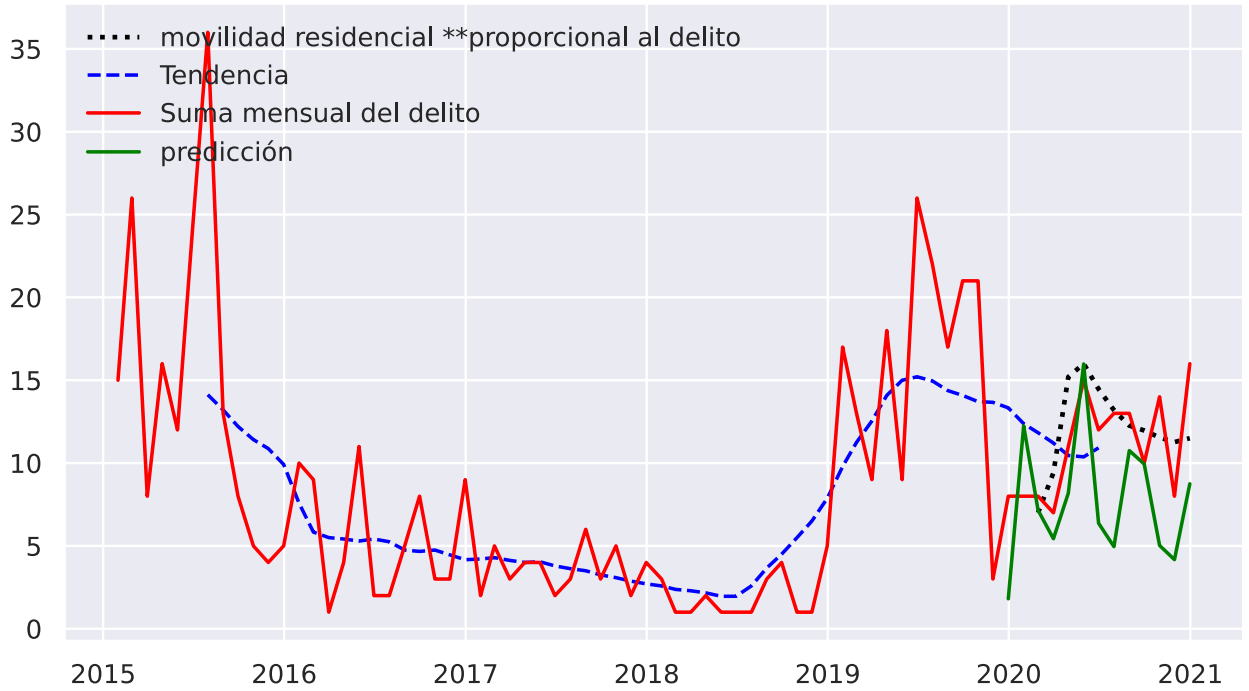
Narcomenudeo en Chihuahua  
 correlación del delito con movilidad: 0.58  
 tendencia del delito respecto al año anterior: 0.05



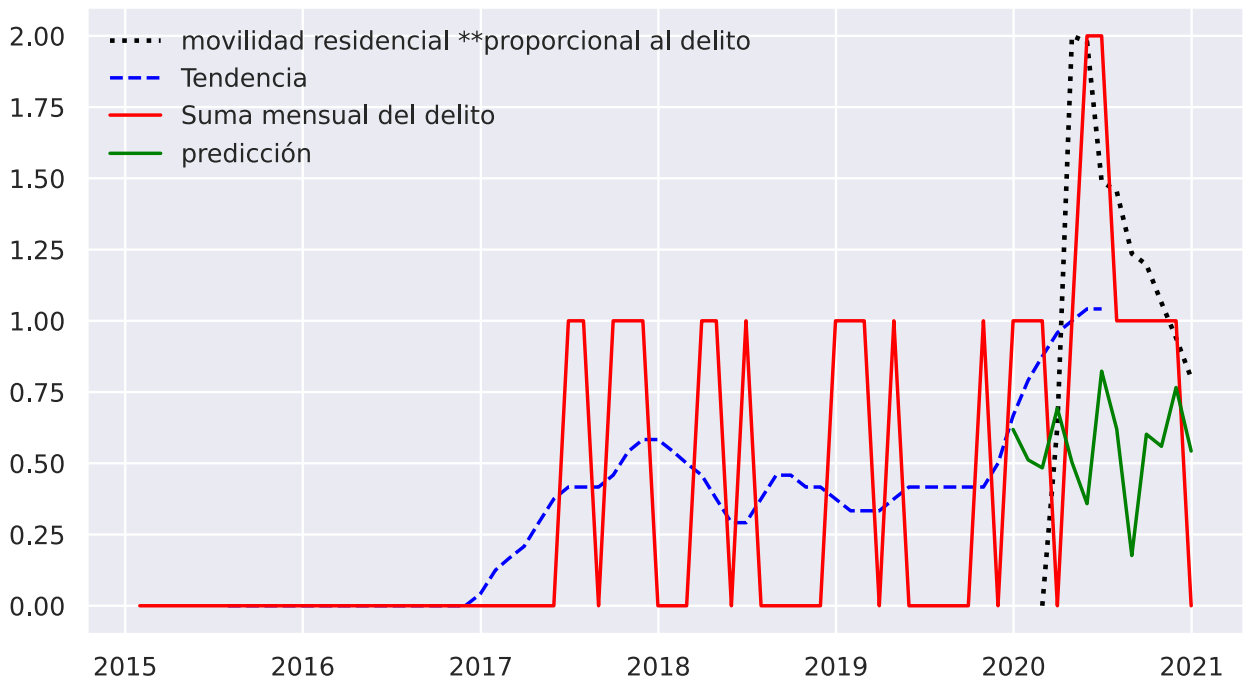
Incesto en Baja California  
 correlación del delito con movilidad: 0.57  
 tendencia del delito respecto al año anterior: -0.67



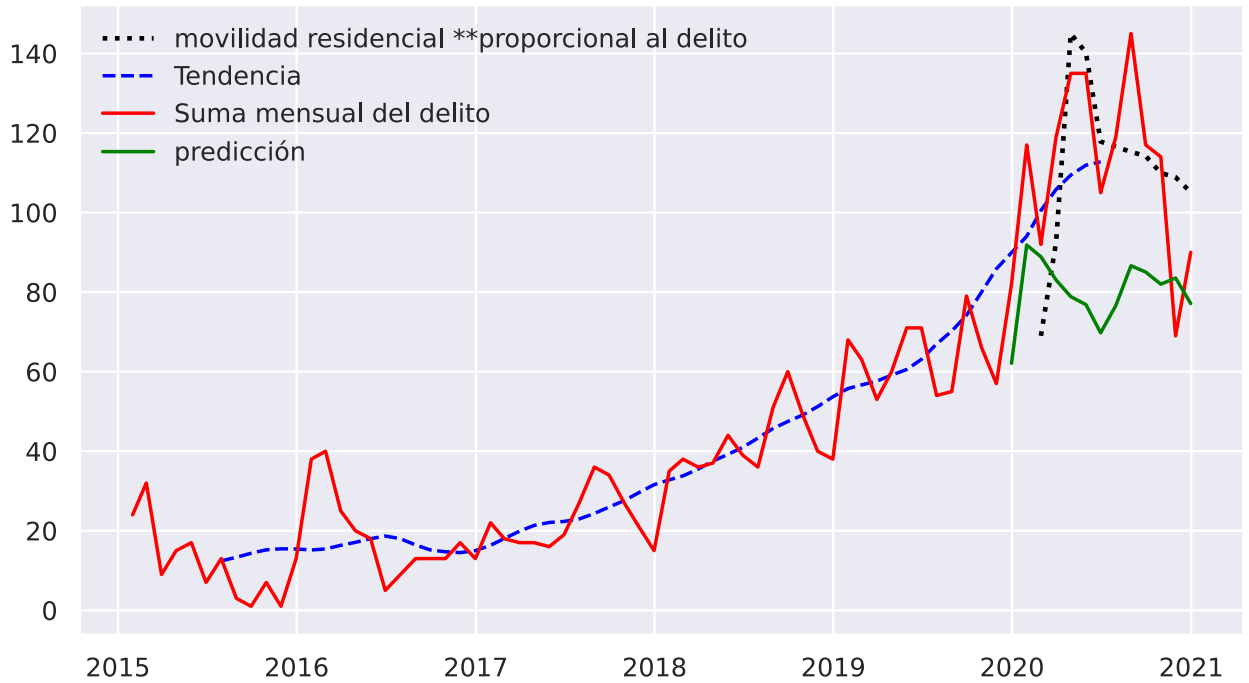
Extorsión en Morelos  
 correlación del delito con movilidad: 0.57  
 tendencia del delito respecto al año anterior: -0.27



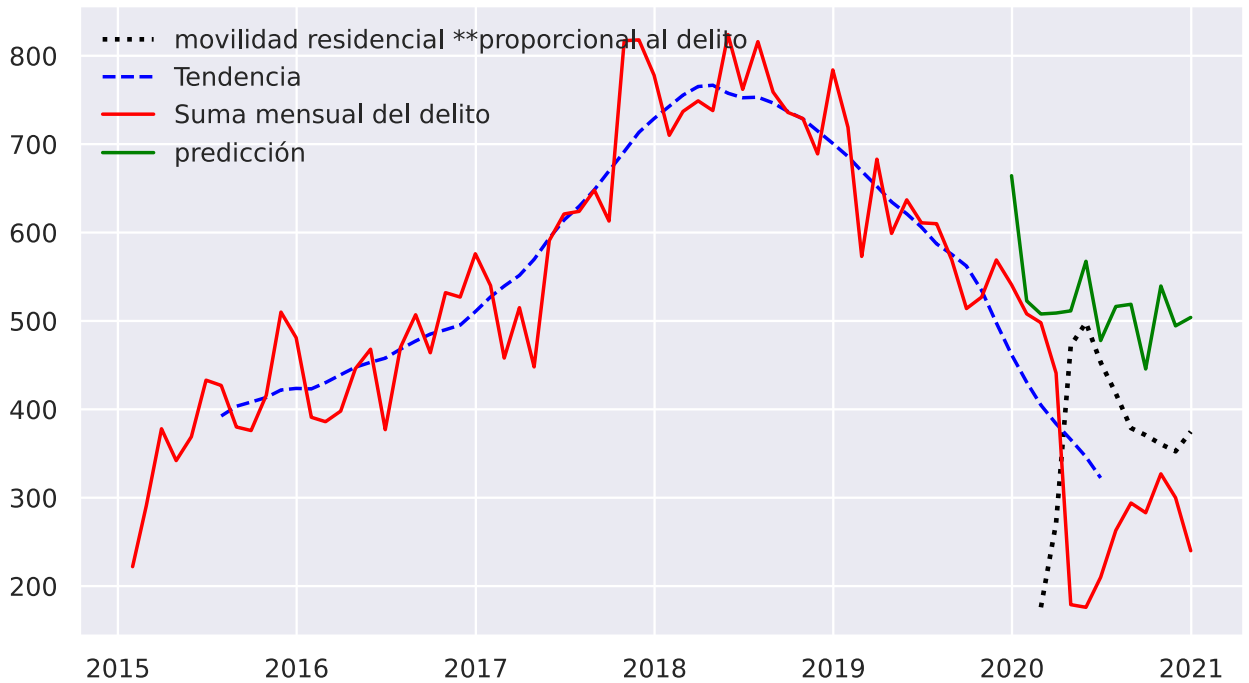
Feminicidio en Nayarit  
 correlación del delito con movilidad: 0.55  
 tendencia del delito respecto al año anterior: 1.4



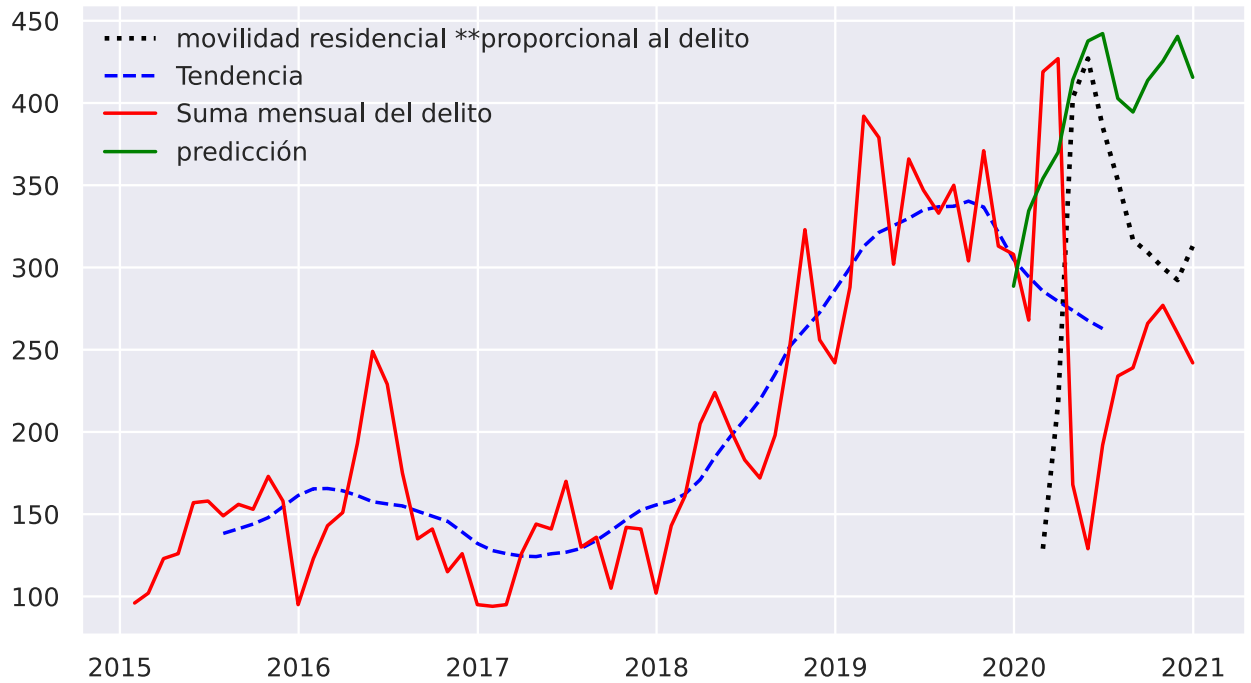
Narcomenudeo en San Luis Potosí  
correlación del delito con movilidad: 0.54  
tendencia del delito respecto al año anterior: 0.74



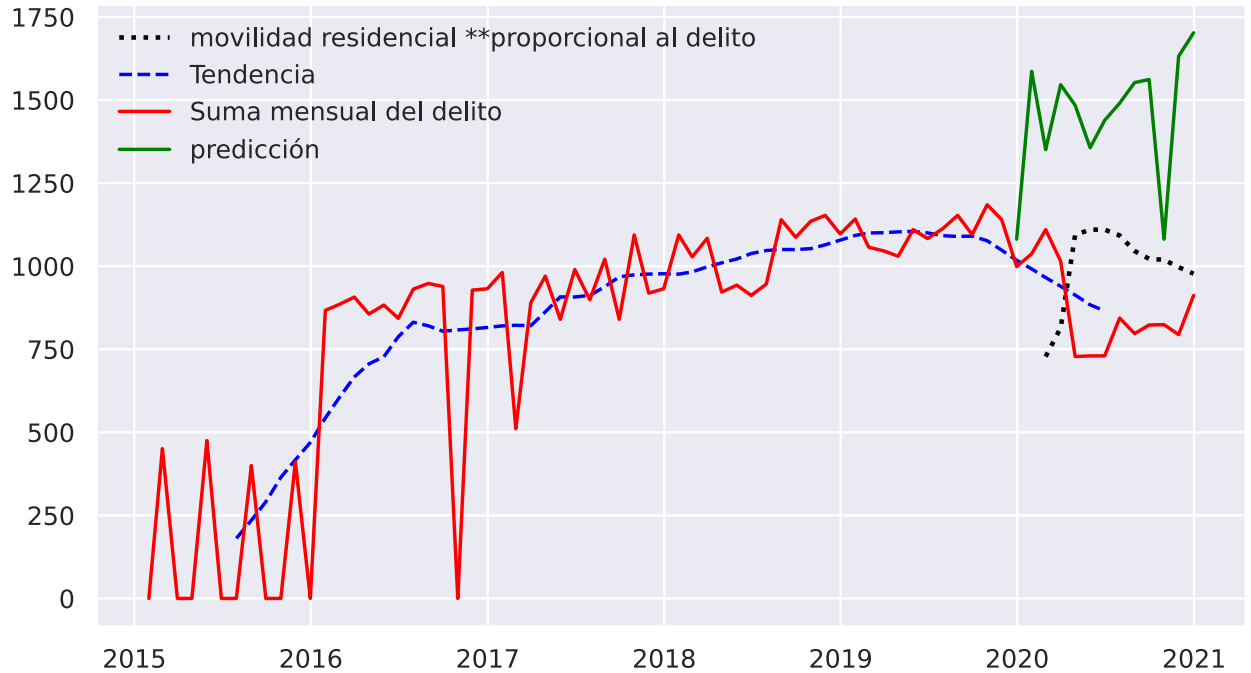
Falsificación en Ciudad de México  
correlación del delito con movilidad: -0.97  
tendencia del delito respecto al año anterior: -0.48

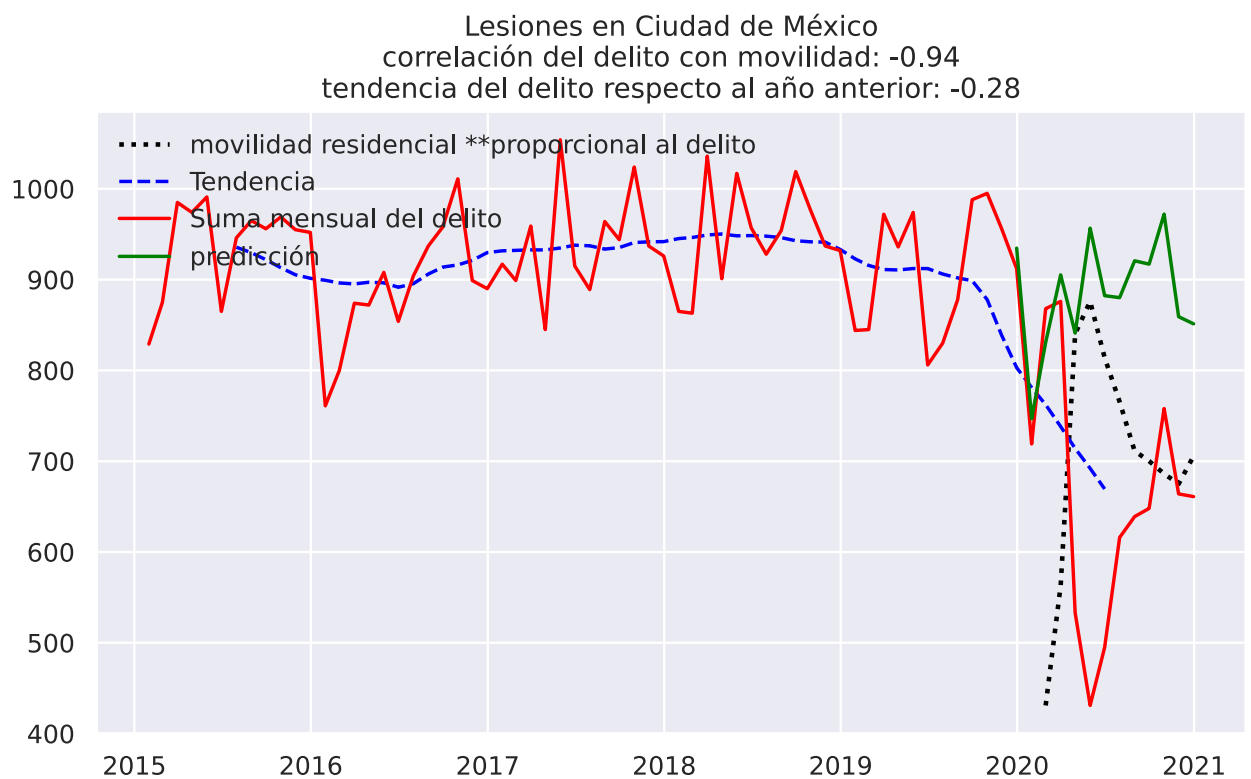
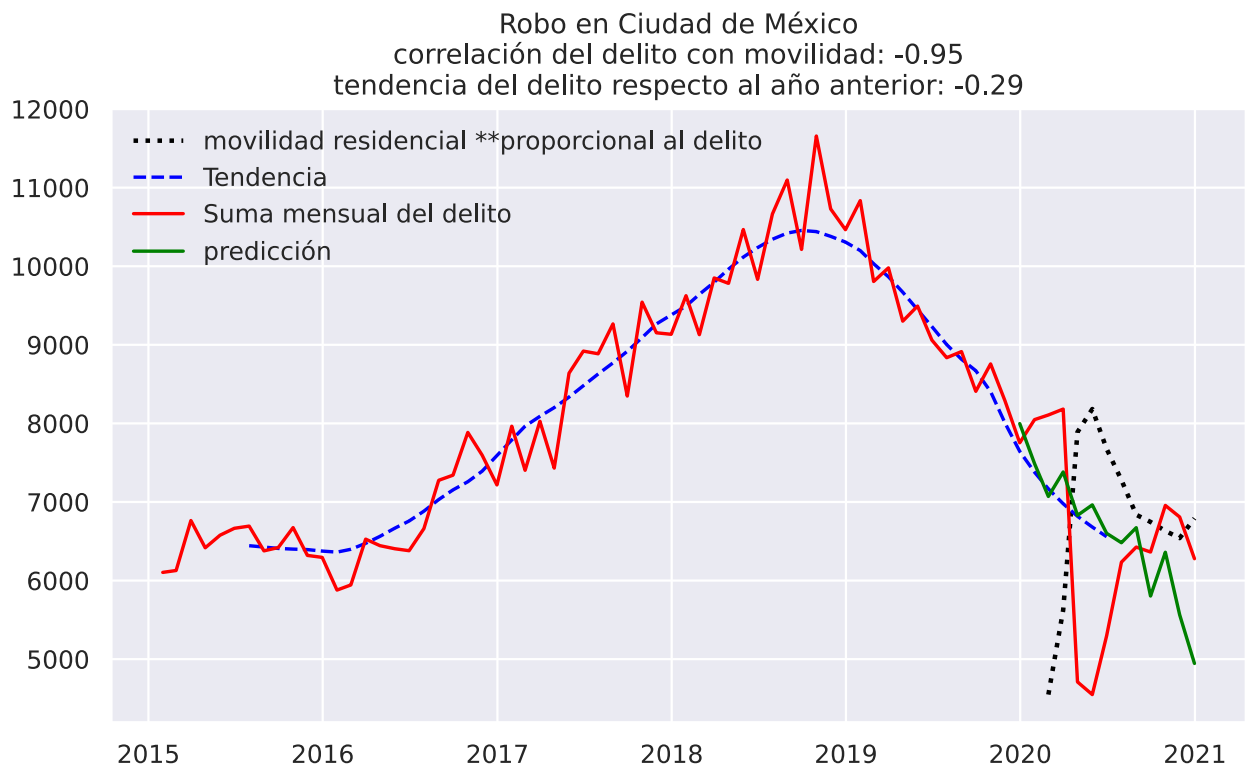


Abuso sexual en Ciudad de México  
 correlación del delito con movilidad: -0.96  
 tendencia del delito respecto al año anterior: -0.23



Robo en Oaxaca  
 correlación del delito con movilidad: -0.95  
 tendencia del delito respecto al año anterior: -0.21





In [ ]: