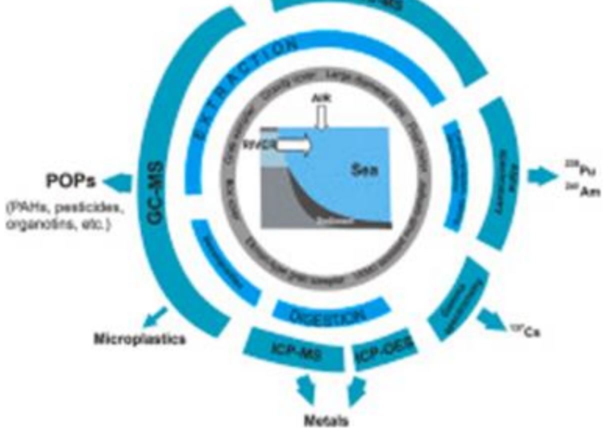
Marine sediment analysis

A review of advanced approaches and practices focused on contaminants

Abstract:

This review is aimed at critical analysis of current and emerging capabilities of analytical methods as employed for marine sediment analysis. An emphasis is given to the most reliable experimental strategies used to quantifying the various classes of contaminants and thus to acquiring analytical information that is relevant to assess the quality of sediments with regard to possible pollution of marine ecosystems. Advanced analytical methodology in use basically relies on the application of mass spectrometry to enable identification and quantification of hazardous chemical species, directly or after separation using the principles of gas and liquid chromatography. Also addressed are sample preparation techniques which – given the complexity of sediment matrices and the diverse and multiple nature of contaminants – are often a key for successful analysis. Among the trends in marine sediment analysis is an inclination to chemicals that are only recently recognized as emerging pollutants of very high concern, such as microplastics, pharmaceuticals and their metabolites.

Graphical abstract:



Introduction:

Analysis of marine sediments is the field of analytical chemistry that is closely associated with the impact of anthropogenic factors on the environment. Sediments are a large storage of various pollutants, especially in regions with intense industrial activities, oil exploration, shipping, etc. When released into the overlying water through physical (erosion, resuspension, deposition), biological (bioturbation) and chemical (desorption and benthic diffusion) processes, the accumulated toxicants exert a strong influence on the ecological health of marine ecosystems. Therefore, sediment surveying in terms of the level of potential contaminants is a major concern to assess the state of the marine environment and to provide baseline guidance for relevant control authorities.

Section snippets:

Organic contaminants of an array of sediment pollutants of organic nature, POPs represent the widest class of chemicals. They are mostly lipophilic, highly toxic and rugged in the environment synthetic compounds, being hence transported over long distances and leading to global pollution. Particularly significant and most commonly identified POPs are organochlorines, such as PCBs, PAHs and pesticides, for which marine sediments act as a bioaccumulation pool.

Pharmaceuticals:

Pharmaceuticals and their metabolites have been increasingly detected in different environmental compartments, including marine sediments. This is due to a growing medical and veterinary employment of pharmaceuticals world-wide. Their inherent biological activity poses high risks to the ecosystem and to human health via contact with contaminated sediments.

Heavy metals:

Owing to extreme toxicity, non-biodegradability and high accumulation potential of heavy metals, their presence in aquatic ecosystems is a matter of serious anxiety. The metal content in marine sediments fluctuates being impacted by various anthropogenic factors, such as industrial waste, transportation, agriculture, etc., or naturally but differently enriched from the rocks outcropping in the source area.

Radionuclides:

There are two major routes how radionuclides find their way into the environment and may then cause harmful effects on living organisms including humans. One is radiation accidents unavoidably occurring in nuclear facilities, i.e. nuclear power and research reactors or involved in nuclear-weapons programs. The second is due to still no safe means for final geological disposal of radioactive waste produced by nuclear reactors.

Microplastics:

Of other classes of relevant contaminants, plastic materials with micro- or nano-sizes have become of increasing apprehension, even though the associated dangers after final disposal at the deep sea-floor and mobilization by natural processes need to be unambiguously confirmed. Such small sizes can be a result of degradation or fragmentation of larger plastic materials upon their waste into the environment or intentional production (and disposal) of manufactured microplastics.

Exploratory Analysis:

To begin this exploratory analysis, first use matplotlib to import libraries and define functions for plotting the data. Depending on the data, not all plots will be made. (Hey, I'm just a kerneling bot, not a Kaggle Competitions Grandmaster.

```
from mpl_toolkits.mplot3d import Axes3D

from sklearn.preprocessing import StandardScaler

import matplotlib.pyplot as plt # plotting

import numpy as np # linear algebra

import os # accessing directory structure

import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
```

There is 1 csv file in the current version of the dataset:

```
print(os.listdir('../input'))
```

The next hidden code cells define functions for plotting data. Click on the "Code" button in the published kernel to reveal the hidden code.

```
# Distribution graphs (histogram/bar graph) of column data

def plotPerColumnDistribution(df, nGraphShown, nGraphPerRow):

        nunique = df.nunique()

        df = df[[col for col in df if nunique[col] > 1 and nunique[col] <50]]   # For displaying purposes, pick columns that have between 1 and 50 unique values

        nRow, nCol = df.shape

        columnNames = list(df)

        nGraphRow = (nCol + nGraphPerRow - 1) / nGraphPerRow

        plt.figure(num = None, figsize = (6 * nGraphPerRow, 8 *nGraphRow), dpi = 80, facecolor = 'w', edgecolor = 'k')

        for i in range(min(nCol, nGraphShown)):

                plt.subplot(nGraphRow, nGraphPerRow, i + 1)

                columnDf = df.iloc[:, i]

                if (not np.issubdtype(type(columnDf.iloc[0]), np.number)):

                        valueCounts = columnDf.value_counts()

                        valueCounts.plot.bar()

                else:

                        columnDf.hist()

                plt.ylabel('counts')

                plt.xticks(rotation = 90)

                plt.title(f'{columnNames[i]} (column {i})')
```

```
        plt.tight_layout(pad = 1.0, w_pad = 1.0, h_pad = 1.0)
        plt.show()
# Correlation matrix
def plotCorrelationMatrix(df, graphWidth):
    filename = df.dataframeName
    df = df.dropna('columns') # drop columns with NaN
    df = df[[col for col in df if df[col].nunique() > 1]]   # keep columns where there are more than
1 unique values
    if df.shape[1] < 2:
        print(f'No correlation plots shown: The number of non-NaN or constant columns
({df.shape[1]}) is less than 2')
        return
    corr = df.corr()
    plt.figure(num=None, figsize=(graphWidth, graphWidth), dpi=80, facecolor='w', edgecolor='k')
    corrMat = plt.matshow(corr, fignum = 1)
    plt.xticks(range(len(corr.columns)), corr.columns, rotation=90)
    plt.yticks(range(len(corr.columns)), corr.columns)
    plt.gca().xaxis.tick_bottom()
    plt.colorbar(corrMat)
    plt.title(f'Correlation Matrix for {filename}', fontsize=15)
    plt.show()
# Scatter and density plots
def plotScatterMatrix(df, plotSize, textSize):
    df = df.select_dtypes(include =[np.number])   # keep only numerical column   # Remove rows
and columns that would lead to df being singular
    df = df.dropna('columns')
    df = df[[col for col in df if df[col].nunique() > 1]] # keep columns where there are more than
1 unique values
    columnNames = list(df)
    if len(columnNames) > 10: # reduce the number of columns for matrix inversion of kernel
density plots
        columnNames = columnNames[:10]
    df = df[columnNames]
```

```
ax = pd.plotting.scatter_matrix(df, alpha=0.75, figsize=[plotSize, plotSize], diagonal='kde')

corrs = df.corr().values

for i, j in zip(*plt.np.triu_indices_from(ax, k = 1)):

        ax[i, j].annotate('Corr. coef = %.3f' % corrs[i, j], (0.8, 0.2), xycoords='axes fraction', ha='center', va='center', size=textSize)

plt.suptitle('Scatter and Density Plot')

plt.show()
```

Now you're ready to read in the data and use the plotting functions to visualize the data.

Let's check 1st file: ../input/Tweets.csv

```
nRowsRead = 1000   # specify 'None' if want to read whole file # Tweets.csv has 14640 rows in reality, but we are only loading/previewing the first 1000 rows

df1 = pd.read_csv('../input/Tweets.csv', delimiter=',', nrows = nRowsRead)

df1.dataframeName = 'Tweets.csv'

nRow, nCol = df1.shape

print(f'There are {nRow} rows and {nCol} columns')
```

Let's take a quick look at what the data looks like:

```
df1.head(5)
```

Distribution graphs (histogram/bar graph) of sampled columns:

```
plotPerColumnDistribution(df1, 10, 5)
```

Conclusions and perspectives:

The literature regarding the analysis of marine sediments continues to expand, numbering over 10000 documents in press over the past decade (as provided by Sci finder database). A great proportion of these contributions are devoted to assessment of the ecological state of various marine environments as based on identification and determination of primary or emerging sedimentary contaminants.