Long-term spatiotemporal variability of whitings in Lake Geneva from multispectral remote sensing and machine learning

Gaël Many[1], Nicolas Escoffier[2], Michele Ferrari[1], Philippe Jacquet[1], Daniel Odermatt[3], Grégoire Mariéthoz[4], Pascal Perolo[1], and Marie-Elodie Perga[5]

[1]Université de Lausanne
[2]University of Lausanne - Institute of Earth Surface Dynamics
[3]Eawag - Swiss Federal Institute of Aquatic Science and Technology
[4]University of Lausanne
[5]Universiy of Lausanne

November 24, 2022

Posted on 24 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10512504.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Abstract

Whiting events are massive calcite precipitation events turning hardwater lake waters to a milky turquoise color. The transitory nature of whitings and their variable spatial extent make them poorly captured by traditional monitoring. Herein, we use a multispectral remote sensing approach to describe the spatial and temporal occurrences of whitings in Lake Geneva from 2013 to 2021. Landsat-8, Sentinel-2, and Sentinel-3 sensors are combined and intercalibrated to derive the AreaBGR index and identify whitings using appropriate filters. 95% of the detected whitings are located in the northeastern part of the lake and occur in a highly reproducible environmental setting: a high Rhone River discharge (358.6 +/- 102.1 m^3 s^{-1}), air and water temperatures of 21.3 +/- 3.0 °C and 18.0 +/- 1.9 °C respectively, and during the stratified period (thermocline depth of 11.1 +/- 0.6 m). An extended time series of whitings in the last 60 years is reconstructed from a random forest algorithm and analyzed through a Bayesian decomposition for annual and seasonal trends in the number of whiting days. Results show that the annual number of whiting days between 1958 and 2021 does not follow any particular monotonic trend. The inter-annual changes of whiting occurrences significantly correlate to the Western Mediterranean Oscillation Index (WeMOI). Besides, spring whitings have increased since 2000 and significantly follow the Atlantic Multidecadal Oscillation index (AMO). Future climate change in the Mediterranean Sea and the Atlantic Ocean could induce more variable and earlier whiting events in Lake Geneva.

Decision Trees and Random Forests

To run the following analysis we used Jupyter Notebook with Python 3.10 and several libraries shown below. We also used the Graphitz software (https://graphviz.org/download/).

We will need the following libraries:

In [1]:
```
from sklearn.tree import DecisionTreeClassifier,export_graphviz
from sklearn.metrics import confusion_matrix,roc_curve,roc_auc_score
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from matplotlib.colors import ListedColormap
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import numpy as np
import matplotlib
import sklearn
import graphviz
import random

import os
os.environ["PATH"] += os.pathsep + 'C:/Users/gaelm/DT_and_RF/Graphviz/bin/'
```

In [2]:
```
from platform import python_version
print("Python:",python_version())
print("sklearn:",sklearn.__version__)
print("matplotlib:",matplotlib.__version__)
print("seaborn:",sns.__version__)
print("pandas:",pd.__version__)
print("numpy:",np.__version__)
print("graphviz:",graphviz.__version__)
```

```
Python: 3.9.12
sklearn: 1.0.2
matplotlib: 3.5.1
seaborn: 0.11.2
pandas: 1.4.2
numpy: 1.21.5
graphviz: 0.20
```

Exploration

In [3]:
```
%matplotlib inline
```

Let us have a look at the dataset:

In [4]:
```
data=pd.read_table("observations_class.txt",sep=',',header=None,names=['date','Rhône disch.','Air Temp.','Water Temp.','Wind Spd.','Thermo. dph','Prec.','class'])
data.drop('date',inplace=True,axis=1)
data.head(3)
```

Out[4]:

	Rhône disch.	Air Temp.	Water Temp.	Wind Spd.	Thermo. dph	Prec.	class
0	174.410	15.4	8.584312	1.4	12	1.4	0
1	183.025	11.1	10.004646	1.5	14	0.0	0
2	228.331	16.4	11.574281	1.6	12	0.0	0

In [5]:
```
data[data.isna().any(axis=1)]
```

Out[5]:

Rhône disch.	Air Temp.	Water Temp.	Wind Spd.	Thermo. dph	Prec.	class

In [6]:
```
data=data.dropna()
```

In [7]:
```
data.shape
```

Out[7]:
```
(861, 7)
```

There are 861 data points with 6 input features and 1 output feature.

In [8]:
```
data.describe()
```

Out[8]:

	Rhône disch.	Air Temp.	Water Temp.	Wind Spd.	Thermo. dph	Prec.	class
count	861.000000	861.000000	861.000000	861.000000	861.000000	861.000000	861.000000
mean	209.576177	14.592451	14.323904	2.311034	25.189315	0.709059	0.123113
std	106.310423	7.268774	4.937211	1.407333	42.346386	2.855413	0.328757
min	57.972000	-6.800000	6.299961	0.800000	11.000000	0.000000	0.000000
25%	124.269000	8.900000	9.210113	1.400000	11.000000	0.000000	0.000000
50%	176.643000	15.800000	14.916334	1.900000	12.000000	0.000000	0.000000
75%	281.508000	20.600000	19.127007	2.600000	17.000000	0.000000	0.000000
max	617.673000	28.700000	22.893131	10.100000	310.000000	29.900000	1.000000

In [9]:
```
data.groupby("class").size()
```

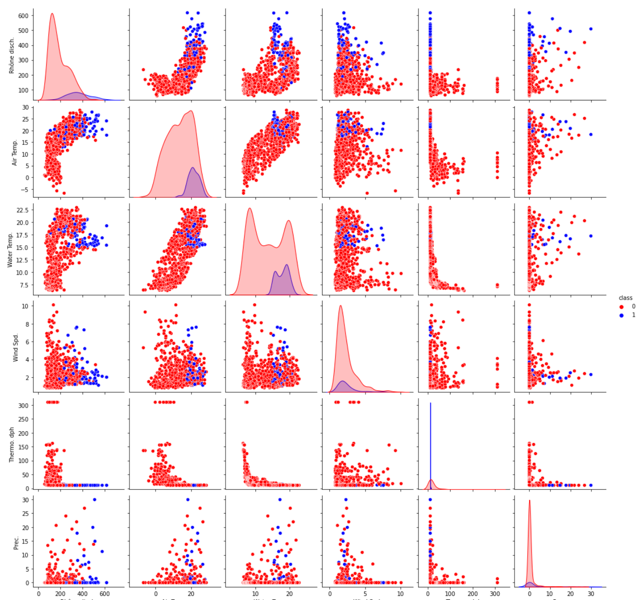
Out[9]:
```
class
0    755
1    106
dtype: int64
```

The dataset is quite imbalanced, so we will use weights in the models to make the "rare" class comparable to the "normal" class.

Let us look at the pair plots:

In [10]:
```
colors=["red","blue"]
customPalette=sns.set_palette(sns.color_palette(colors))
sns.pairplot(data,hue='class',palette=customPalette)
```

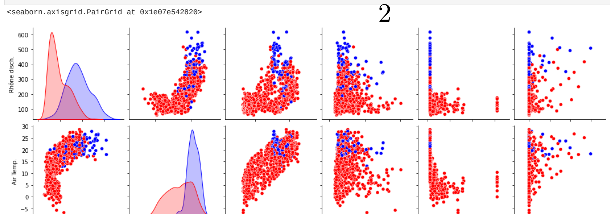
Out[10]:
```
<seaborn.axisgrid.PairGrid at 0x1e07b58b730>
```



Let us use the option "common_norm=False" to normalize separately the red and blue classes:

In [11]:
```
sns.pairplot(data,hue='class',palette=customPalette,diag_kws=dict(common_norm=False))
```

Out[11]:
```
<seaborn.axisgrid.PairGrid at 0x1e07e542820>
```

Posted on 24 Nov 2022 — CC-BY 4.0 — https://doi.org/10.1002/essoar.10512504.1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Long-term spatiotemporal variability of whitings in Lake Geneva from multispectral remote sensing and machine learning

G. Many[1*], N. Escoffier[1], M. Ferrari[1], P. Jacquet[2], D. Odermatt[3], G. Mariethoz[1], P. Perolo[1], M-E. Perga[1]

[1] Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

[2] Scientific Computing and Research Support Unit, University of Lausanne, Lausanne, Switzerland

[3] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Surface Waters, Research and Management, Dübendorf, Switzerland

Corresponding author: Gaël Many (gael.many@unil.ch)

Key Points:

- Machine learning and remote sensing techniques are combined to study whitings in Lake Geneva from 1958 to 2021
- The Western Mediterranean oscillation explains inter-annual changes in whiting occurrences
- The Atlantic multi-decadal oscillation has been responsible for earlier whitings since 2000

Abstract

Whiting events are massive calcite precipitation events turning hardwater lake waters to a milky turquoise color. The transitory nature of whitings and their variable spatial extent make them poorly captured by traditional monitoring. Herein, we use a multispectral remote sensing approach to describe the spatial and temporal occurrences of whitings in Lake Geneva from 2013 to 2021. Landsat-8, Sentinel-2, and Sentinel-3 sensors are combined and intercalibrated to derive the AreaBGR index and identify whitings using appropriate filters. 95% of the detected whitings are located in the northeastern part of the lake and occur in a highly reproducible environmental setting: a high Rhone River discharge (358.6 +/- 102.1 $m^3 s^{-1}$), air and water temperatures of 21.3 +/- 3.0 °C and 18.0 +/- 1.9 °C respectively, and during the stratified period (thermocline depth of 11.1 +/- 0.6 m). An extended time series of whitings in the last 60 years is reconstructed from a random forest algorithm and analyzed through a Bayesian decomposition for annual and seasonal trends in the number of whiting days. Results show that the annual number of whiting days between 1958 and 2021 does not follow any particular monotonic trend. The inter-annual changes of whiting occurrences significantly correlate to the Western Mediterranean Oscillation Index (WeMOI). Besides, spring whitings have increased since 2000 and significantly follow the Atlantic Multidecadal Oscillation index (AMO). Future climate change in the Mediterranean Sea and the Atlantic Ocean could induce more variable and earlier whiting events in Lake Geneva.

1 Introduction

Calcium carbonate precipitation is an essential biogeochemical process in freshwater and marine ecosystems (Ridgwell and Zeebe, 2005; Khan et al., 2022). In hardwater lakes, calcite precipitation represents a major component of the inorganic carbon cycle. Calcite precipitation also interferes with lake nutrient cycles owing to its complexation with phosphates (Müller et al., 2016). Calcite precipitation is a seasonal process that can occur discreetly at a low background level. However, under favorable conditions, it can also manifest more strikingly through massive short-term transitory events, so-called whiting events. Whiting events are common phenomena of marine environments (Sondi and Juracic, 2010; Larson and Mylroie, 2014; Shanableh et al., 2019; 2021) and lakes (Strong and Eadie, 1978; Effler, 1987; Thompson et al., 1997; Nouchi et al., 2019). Whitings are characterized by a milky turquoise coloration of upper surface layers, generated by a fine-grained size of calcite precipitates that increase the turbidity of the water column and its light reflectance (Peng and Effler, 2017).

The supersaturation of surface waters for calcite is a necessary but insufficient prerequisite for mineral precipitation and thus whiting events. Calcite supersaturation can be reached through a shift in carbonate equilibria induced by an increase in pH or CO_2 removal (Müller et al., 2016) along with greater water temperatures that decrease the retrograde solubility of calcite (Hodell et al., 1998). However, homogeneous nucleation requires overpassing the activation energy far above the strict supersaturation. Massive events such as whitings require adequate nucleation seeds for heterogeneous precipitation in the water column (Stabel, 1986). In hardwater lakes, whiting events have mainly been associated with phytoplankton activity. For instance, picoplankton growth can create the requested pH and CO_2 conditions for supersaturation, while the cells can act as heterogeneous nuclei (Thompson et al., 1997; Peng and Effler, 2011). Once supersaturation is reached, river-borne detrital particles can also trigger nucleation (Effler and Peng, 2012; Nouchi et al., 2019; Escoffier et al., 2022). Altogether, these observations evidenced that warmer surface temperatures, enhanced primary production, and fine suspended sediments can potentially all contribute to whiting events, even though their interplay may vary from one lacustrine system to another. Moreover, whiting events are likely regulated by a broader combination of climatic and trophic factors that are both dynamic in time. Therefore, determining the long-term evolution of whiting events' occurrences in relation to global change impacts on environmental factors (e.g., physical conditions of lakes, changes in river inputs, lakes' primary production) appears crucial for predicting changes in the inorganic carbon cycle of inland waters.

Due to their episodic and transient nature, the dynamics of whiting events can only be captured by high-frequency monitoring. However, whiting events are also patchy in space and can be missed by moored high-frequency sensors. In fact, as the typical turquoise coloration of whiting events usually covers large areas, these phenomena are excellent candidates for remote sensing detection. Whiting events have, for instance, been monitored through remote sensing techniques in diverse marine areas such as the Arabian Gulf (Shanableh et al., 2019; 2021), the Bahamas sea (Dierssen et al., 2009) or Florida coastal waters (Long et al., 2017) as well as in diverse lacustrine systems in Germany (Heine et al., 2017), Switzerland (Nouchi et al., 2019) or North America (Binding et al., 2015). However, while these approaches provided detailed information on the spatial extent of whiting events, they were also characterized by specific limitations in terms of temporal coverage. For instance, remote sensing datasets can be discontinuous due to both the satellite time resolution and a potential absence or limited quality of images associated with cloud cover. Hence, because of this limitation and the restricted availability of time-resolved, multi-annual ground monitoring data, there are few references of continuous records of whiting occurrences long enough to evaluate how their dynamics respond to changing environmental and climatic conditions. For instance, Long et al. (2018) investigated the annual mean whiting occurrence frequency and spatial distribution from MODIS data on a decadal timescale in the coastal waters of Florida. However, they could not provide insights on the underlying drivers. Similarly, Binding et al. (2015) provided an extensive description of water clarity-inferred whiting event dynamics in the Great Lakes on multi-decadal scales. However, they only related the observed changes to reported long-term biogeochemical evolution of the lacustrine systems without statistically exploring the environmental drivers supporting the triggering of whiting events in the short term nor the response of these factors to long-term climatic forcing.

Herein, we aim to use machine learning techniques to combine ground-based and remote sensing data to reconstruct the dynamics of whiting events in a large peri-alpine hardwater lake, Lake Geneva. Accordingly, (i) we use the multispectral long-term remote sensing data of Landsat-8, Sentinel-2, and Sentinel-3, to determine the spatial and temporal occurrences of whiting events in Lake Geneva from 2013 to 2021. Then, (ii) we apply a random forest machine learning approach to identify, from ground-based monitoring data, the environmental setting during whitings in the lake and reconstruct the past “unseen” whiting days. Finally, (iii) we analyze the temporal dynamics of whiting occurrence over the past 60 years in relation to the relevant climate indices affecting Central Europe.

2 Study site

Lake Geneva is a peri-alpine lake along the Swiss-French border, at 372 meters above sea level (46°26′ N, 6°33′ E, see Figure 1). The lake's surface area is about 580 km^2, and its maximum depth (309 m) makes it the largest freshwater body in Western Europe, with a volume of 89 km^3. Lake Geneva is oligomictic; however, complete mixing happens only during exceptionally cold winters, and recent studies describe the lake as meromictic (Schwefel et al., 2016). On an interannual scale, the long-term variability of the Atlantic climate influences the thermal conditions of Lake Geneva. Subtropical Atlantic activity, reflected by the Atlantic Multidecadal Oscillation (AMO), has been described as the main factor influencing summer conditions in the lake (Molinero et al., 2007). Winter conditions have been mostly related to the activity of the North Atlantic, reflected by the Northern Atlantic Oscillation (NAO, Ottersen et al., 2001).

The main tributary to Lake Geneva is the Rhone, representing approximately 70% of the total water input. The Rhone River is also the primary supplier of sediment and phosphate to the lake (Loizeau and Dominik, 2000; Perga et al., 2016) and plays a major role in lake ecosystem dynamics in terms of biogeochemical processes (primary production, fine sediments delivery, transport, and settling; Lambert and Perga, 2019; Escoffier et al., 2022). On the interannual scale, rainfall and summer temperature changes are expected to play a role in discharge variability. The Atlantic (AMO, NAO), Mediterranean (such as Western Mediterranean Oscillation Index; WeMOi), and even global (Oceanic Nino Index; ONI) climate indices appear to be crucial in describing this variability.

The inflowing water from the Rhône generally takes the form of an interflow when the lake is thermally stratified, i.e., a turbid layer that propagates along the thermocline where the Rhône water finds its neutral buoyancy (Giovanolli, 1990). However, these particulate inputs can also flow along the bottom of the lake when extreme densities are reached (cold water and high concentration of suspended particles). During these events, the Rhone inflow is not observable by satellite. However, extreme discharge events when the lake is not stratified can cause overflows detraining suspended particles toward surface waters. These events, episodically visible by remote sensing, are poorly described in the literature. It is therefore important to discriminate these events from whiting events in Lake Geneva, which will be addressed in this study.

Recent studies on whiting events in Lake Geneva have been carried out by in situ measurements, remote sensing, and hydrological modeling. So far, whiting events have been observed in late spring/early summer when 1) the Rhône discharge is high due to catchment snowmelt, and 2) the lake's waters are stratified and surface temperatures are warm. Nouchi et al. (2019) demonstrated that whiting events are triggered along the Rhône interflow into the lake and that its spatial extent, influenced by local hydrodynamics, corresponds to the northeastern dispersion of riverine particles. Besides, Escoffier et al. (2022) filled in the gap of in situ monitoring of whiting dynamics. They showed that there are different contributions of in situ $CaCO_3$ particles. A detrital part eroded from the Rhône catchment and brought into the interflow, and an authigenic part (i.e., newly formed $CaCO_3$ particles), probably precipitated on the surface of fine fluvial particles transported into the lake. This authigenic calcite component tends to increase with distance from the mouth of the Rhône, highlighting the role of the physical stability of the water column and the spread of the interflow in the dynamics of whitings in Lake Geneva.

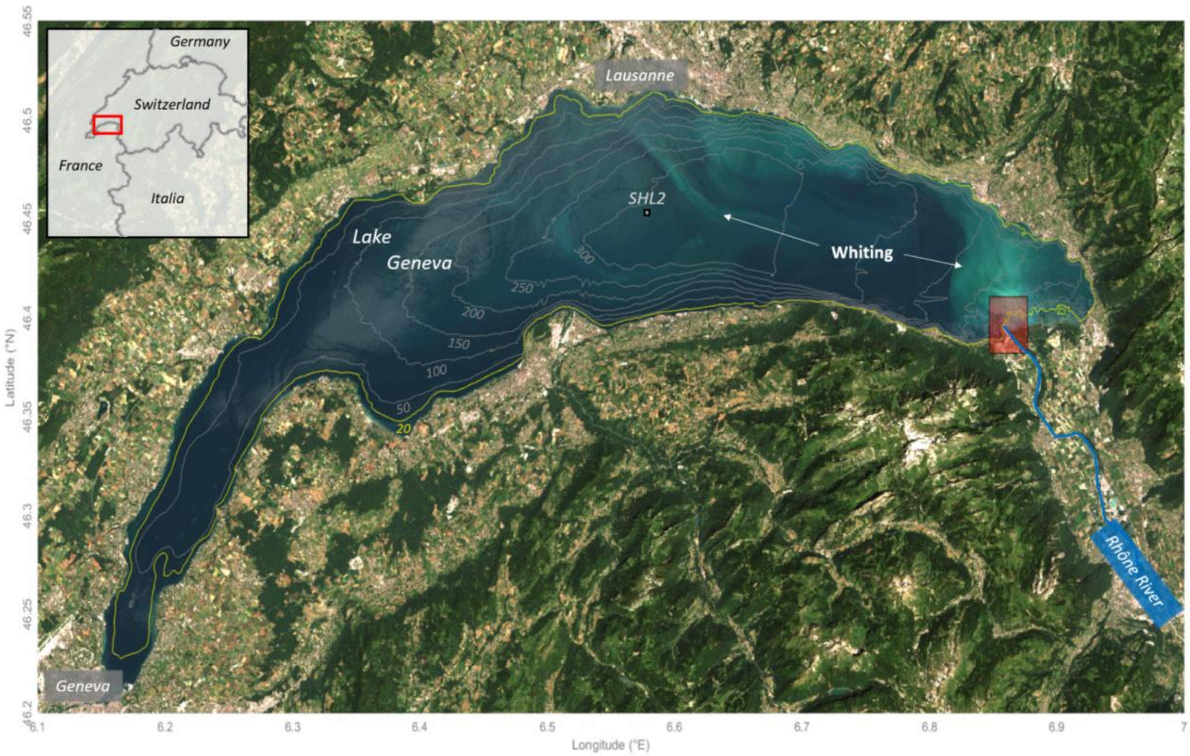


Figure 1. Map of the study area. RGB image from Landsat-8 of Lake Geneva on 29 June 2019. The whiting areas (i.e., turquoise 'milky' color of surface waters) are specified. The SHL2 monitoring point is shown in grey in the middle of the lake. The Rhone River is shown in blue. The lake's location between France and Switzerland is shown in the top-left corner. The 20m isobath is shown in yellow and the Rhone estuary area in red.

3 Workflow and data

3.1 Workflow

The workflow consists of multiple processing steps from remote sensing images selection, data filtering (region of interest, 30% cloud cover filtering), whiting index estimates, and data export using the cloud computing platform Google Earth Engine (GEE) (Kumar and Mutanga, 2018; Mutanga and Kumar, 2019) for Landsat-8 and Sentinel-2 data and from Datalakes (https://www.datalakes-eawag.ch/) for Sentinel-3 data. The next processing steps are computed in Matlab. They comprise a sensor response inter-calibration and identify and characterize whiting events. The final, aggregated metrics include the spatial extent and temporal occurrence of whiting events. Factors controlling whiting events in 2013 to 2021 are then studied through decision tree and random forest algorithms, computed in Python. Next, whiting events are classified using environmental indicators, such as meteorological data, Rhone River discharge, and the lake physical conditions. Finally, the optimized random forest is used to reconstruct 'unseen' whiting days from 1958 to 2021.

3.2 Satellite data

Landsat-8, Sentinel-2, and Sentinel-3 satellites are considered in this work. Landsat-8 satellite has a 16-day temporal resolution (under cloud-free conditions; see Table 1 for details).

Landsat-8 carries the Operational Land Imager (OLI), which collects image data in nine visible to shortwave infrared bands with a spatial resolution of 30m. We use the Landsat-8 Collection 1 Tier 1 Raw Scenes (L1TP) provided by USGS on GEE platform to produce the reflectance factors in the RGB bands (Wulder et al., 2019).

The Copernicus Sentinel-2 mission comprises two satellites. The satellites' Multispectral Imager (MSI) acquires data in high temporal resolution (5 days with two satellites at the equator under cloud-free conditions), high spatial resolution (10-60 m pixels, swath width of 290km) and 13 spectral bands ranging from visible to shortwave infrared wavelengths. Sentinel-2 Level-2A data are available on GEE platform. Data are downloaded from the Copernicus datahub and are processed using sen2cor to produce the reflectance factors in the RGB bands (Muller-Wilm et al., 2013). Finally, images are exported from GEE using a spatial resolution of 30m to correspond to the Landsat-8 dataset.

Sentinel-3 satellites (3A and 3B) have a daily temporal resolution. They carry the Ocean and Land Colour Instrument (OLCI), which acquires data along 21 spectral bands ranging from visible to shortwave infrared wavelengths. Medium-resolution (300m) images are processed using the Python package SenCast (https://gitlab.com/eawag-rs/sencast). Normalized water-leaving reflectance in the RGB bands is calculated using the Polymer algorithm v4.13 (Steinmetz and Ramon, 2018), which is tried and tested for lake water quality retrieval in the Copernicus Global Land Service (Copernicus, 2020) and ESA's Climate Change Initiative (ESA, 2020). All Sentinel-3 data used in this study are available in the Datalakes webportal (www.datalakes-eawag.ch).

Table 1. Specifications of the Landsat-8 OLI, Sentinel-2 MSI and Sentinel-3 OLCI data used for the study. The number of cloud-free images available during the period of interest is specified. *Nominal temporal resolution. Actual temporal resolution depends on the cloudiness of the study area. ** MSI data was resampled to 30m to fit with the resolution of OLI data.

Sensor	OLI	MSI	OLCI
Spatial resolution (m)	30	10-60**	300
Swath width (km)	180	290	1270
Temporal resolution* (days)	16	5	1
Available period	2013-2021	2017-2021	2016-2021
λ_{blue}	480	490	490
λ_{green}	560	560	560
λ_{ref}	655	665	665
Cloud-free images used	140	101	766

3.3 Meteorological, monitoring, and climate data

Daily mean meteorological conditions from 1958 to 2021 are downloaded from the MeteoSwiss IDAWEB website (https://gate.meteoswiss.ch/idaweb/login.do). Air temperature and wind speed are measured at the Geneva-Cointrin Station (code station GVE; 6°08'E; 46°15'N). Water temperature profiles measured fortnightly since 1958 are extracted from the SI OLA database (Rimet et al., 2020). Data are interpolated within a 1 m vertical 1-day temporal resolution grid. In this work, surface water temperature (0-10m) is used as a filter to discard false-positive whiting days (see 4.1). The thermocline depth is computed over the entire period (i.e., 1958-2021). Historical discharge data of the Rhone River (1958-2021) are downloaded from the FOEN website (FOEN, 2022). Discharge data are monitored at the Porte du Scex station with a daily resolution.

The climatic indexes tested encompass the AMO (https://www.psl.noaa.gov/data/timeseries/AMO/), which is referenced as a good indicator of the summer climate in central Europe (Molinero et al., 2007), and the NAO (https://www.ncei.noaa.gov/access/monitoring/nao/), which has been described as the main winter climate forcing (Ottersen et al., 2001). Besides, we also test the WeMOi (https://crudata.uea.ac.uk/cru/data/moi/), estimated from the difference between atmospheric pressure from northern Italy to southwestern Spain (Izquierdo et al., 2014). It is representative of rainfall variability in both areas. Positive phases typically show an anticyclone in the Gulf of Cadiz and a low-pressure area over the Ligurian Sea, leading to increased precipitations in northern Italy, and probably in our study area (Martin-Vide and Lopez-Bustins, 2006). Finally, the Oceanic Nino Index (ONI, https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php) is also tested, as this index is referenced as the primary index for tracking El Nino Southern Oscillation phenomenon, which is a major contributor of worldwide climate variability (McPhaden et al., 2006), and potentially a predictable signal in European rainfall (Lloyd-Hughes and Saunders, 2002).

4 Methods

4.1 Whiting detection using remote sensing

The AreaBGR index (see detail in Heine et al., 2017), i.e., the triangular area between the blue, green, and red reflectance values, determines the whiting spatial and temporal occurrences. We use this index as it is the best indicator available to study whiting events in inland waters. The AreaBGR index is computed for all pixels in the abovementioned satellite data of Lake Geneva, using the following expression:

$$\text{AreaBGR} = 0.5\,(\lambda_{blue} * \text{Ref}_{green} + \lambda_{green} * \text{Ref}_{red} + \lambda_{red} * \text{Ref}_{blue} - \lambda_{green} * \text{Ref}_{blue} - \lambda_{red} * \text{Ref}_{green} - \lambda_{blue} * \text{Ref}_{red})$$

An inter-calibration of the different satellite sensors is performed. We compare the AreaBGR estimates for the whiting day on 29 June 2019 for which we have simultaneous images from Landsat-8, Sentinel-2, and Sentinel-3 satellites (see Fig. 2) and ground data (Escoffier et al., 2022). The range of the index measured by Sentinel-2 and Sentinel-3 is slightly lower than that of Landsat-8, as a likely result of different product types and sources, and atmospheric corrections (Heine et al., pers. comm.). The obtained equation $\text{AreaBGR}_{S2} = 0.28 * \text{AreaBGR}_{L8} + 11066.43$ with $R^2 = 0.97$ and $\text{AreaBGR}_{S3} = 0.23 * \text{AreaBGR}_{L8} + 10650.23$ with $R^2 = 0.97$ allows expression of the Sentinel-2 and Sentinel-3 derived AreaBGR indexes in the same range as the one determined by the Landsat-8 satellite (see Fig. 2a). The residuals from the inter-calibration equation can be explained by differences in the sensors' spectral response functions and by the time difference between the shots. Nevertheless, this complementarity allows us to use the Landsat-8 (n=140), Sentinel-2 (n=101), and Sentinel-3 (n=766) databases to describe the spatial and temporal occurrences of whiting days between 2013 and 2021.

Positive whiting is attributed to any pixel whose AreaBGR value is > 13 000, according to Heine et al. (2017) (see magenta contours in Fig. 2b). The surface area of whitings for each image is then estimated by summing flagged pixels of $30m^2$. This database is completed with the daily Sentinel-3 database, from which the AreaBGR is derived following a similar processing. Summing flagged pixels of $300m^2$ provides the area of whiting events.

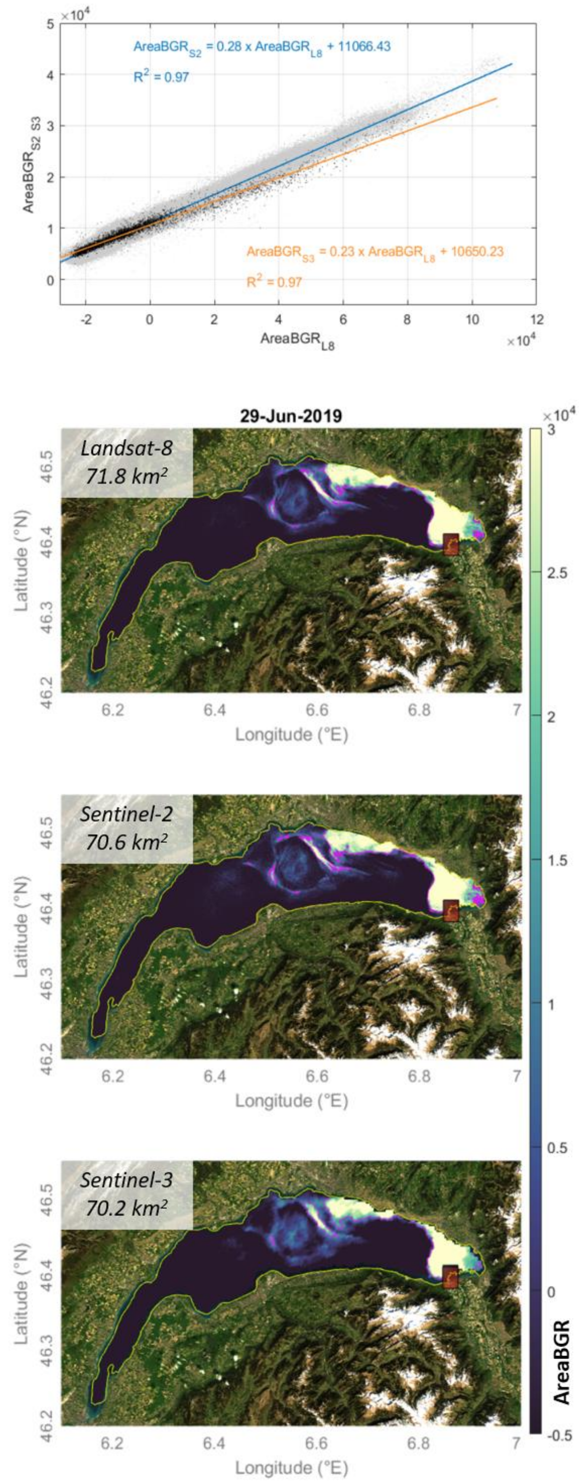


Figure 2. a) Inter-calibration of the AreaBGR index of Landsat-8, Sentinel-2 and Sentinel-3 images on 29 June 2019. Regression between Landsat-8 and Sentinel-2 AreaBGR estimates (grey points) is shown in blue. Regression between Landsat-8 and Sentinel-3 AreaBGR estimates (black points) is shown in orange. The linear equations and correlation coefficients are specified. b) Results of the inter-calibration of AreaBGR of Lake Geneva on 29 June 2019 for Landsat-8,

Sentinel-2, and Sentinel-3 satellites. The whiting area is specified on each image. The delimitation of whiting areas based on the threshold of AreaBGR=13000 is shown in magenta. The 20m isobath used to discard shallow depth is shown in yellow. The Rhone estuary area is shown in red and is discarded from the calculation.

The AreaBGR index can be sensitive to the presence of other suspended particles (Shanableh et al., 2019). In Lake Geneva, in the case of wave-induced resuspension of fine sediments near the coast, AreaBGR may respond to an increase in the near-infrared wavelengths. Events when sediments brought by the Rhone reach the surface (i.e., unstratified lake and cold surface waters, see an example on 25 April 2013 in Fig. 3) generate similar signals. Due to these processes, we apply several filters to discard satellite images showing false-positive whiting days.

First, we only select images with whitings larger than 15km^2 to avoid minor contaminations due to remaining clouds. Then, we exclude the shallowest depths of the lake (i.e., < 20m depth) and the region of the Rhone mouth for our calculations (see the yellow isobath and red area in Fig. 1). Another filter is applied to discard false-positive AreaBGR images due to Rhone inflow at the surface. We base this latter filter on the surface water temperature of the lake (SHL2 monitoring point). Escoffier et al. (2022) showed that whiting events only happened when the lake's surface temperature reaches a minimum of 15°C. Below 15°C, calcite supersaturation is unlikely, while the lake stratification is not strong enough to allow for a Rhone interflow. Therefore, all images with a positive AreaBGR index but surface temperature below 15°C (averaged over 0-10m depth) are discarded.

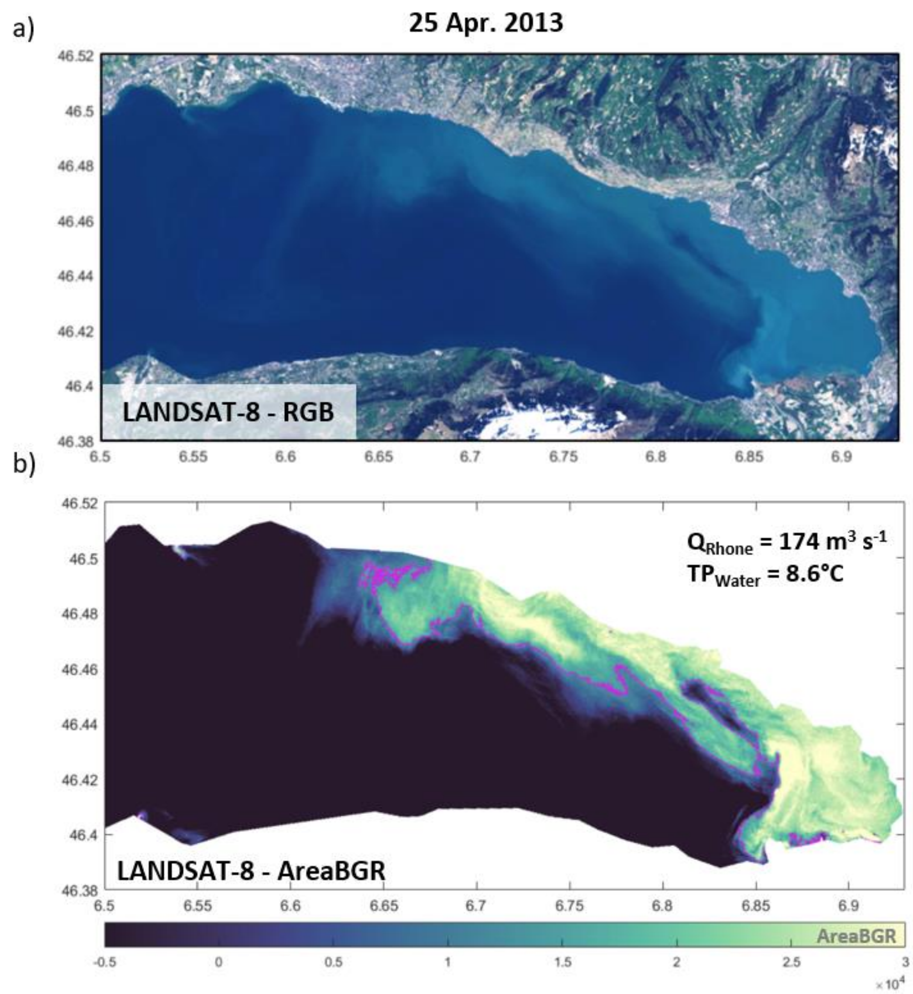


Figure 3. a) Landsat-8 RGB image of Rhone inflow at the lake's surface on 25 April 2013. b) False-positive AreaBGR index caused by the spread of the Rhone inflow at the surface when the lake is unstratified. The area's delimitation based on the area threshold of AreaBGR=13000 is shown in magenta. The Rhone discharge and the lake's surface temperature corresponding to the image's date are specified.

4.2 Reconstruction of past whitings

We use available environmental indicators from 2013 to 2021, i.e. water discharge of the Rhone River, meteorological conditions over Lake Geneva, and the lake physical conditions (surface water temperature, thermocline depth) as input features of a machine learning classification algorithm for whiting occurrence (i.e., whitings or non-whitings, two classes with values of 1 and 0, respectively). The machine learning approach consists of a Decision Tree (DT) and a Random Forest (RF) to find the best classification method based on classical metrics (Hastie et al., 2009; Géron, 2019). The detail of the model development carried out in this work is specified in the Supplementary Material.

First, we split our database into three sub-datasets: (1) the training set (60% of the whole database), (2) the validation set (20%), and (3) the test set (20%). The training set is used to train the different models, i.e., to set the model parameters. The validation set is used to compare the model performances between different models and to choose the most accurate one. The test set is finally used to test the performance of the best model on the remaining 'unused' data.

To evaluate the performances of the models, we use classical metrics such as the confusion matrix (i.e., a table including true negatives, false positives, false negatives, and true positives), the accuracy rate (i.e., the percentage of correct predictions for a given dataset), which is a summary of the confusion matrix, and the AUC (i.e., the Area Under the receiver operating characteristic Curve), which measures how well the whitings and non-whitings events can be separated or distinguished by the model. This Machine Learning approach is expected to provide the main driving factors (among the input features) of the whiting events in Lake Geneva. The best model is then used to reconstruct the past unseen whiting days from 1958 to 2021 relying on the same input features used to train and validate the model for the 2013-2021 period.

Changes in the annual whiting occurrence reconstructed between 1958 and 2021 are tested using Mann-Kendall tests on the time series (Mann, 1945; Kendall, 1948) and a BEAST decomposition (Bayesian Estimator of Abrupt change, Seasonality, and Trend). BEAST is a generic Bayesian model averaging algorithm to decompose time series or 1D sequential data into individual components, such as abrupt changes, trends, and periodic/seasonal variations (Zhao et al., 2019). The relations between the annual whiting frequency and large synoptic climatic indexes are tested using the Pearson correlation coefficient r and the related p-value.

5 Results

5.1 Spatial and temporal occurrences of whitings in Lake Geneva from 2013 to 2021

5.1.1 Spatial occurrences of observed whitings in Lake Geneva

Altogether, 113 whiting days of surface area >15 km^2 are detected in Lake Geneva in 2013-2021. The description of the spatial occurrence of these whiting days, i.e., the number of pixels flagged as whitings between 2013 and 2021, can be challenging as it depends on the available images, i.e., on the temporal resolution and cloud coverage. Note that this result is relative, i.e., a good description of the spatial variability, more than a good estimate of the absolute number of whiting days detected over the study period.

The distribution of whitings by areal coverage is bimodal (Fig. 4a). In 96% of the days, the whiting covers < 40% of the lake area, and exceptional whitings occupy almost the whole lake surface (50-80%). Therefore, we consider them separately (class 1 for partial whitings and 2 for total whitings). Figure 4b-c shows the spatial occurrence of both classes of whiting days as seen by Landsat-8 (2013-2021), Sentinel-2 (2017-2021), and Sentinel-3 (2016-2021). Class 1 whitings are invariably located to the East, near the mouth of the Rhone (class 1; Fig. 4b). Class 2 whitings cover the central part of the lake, even up to the small lake basin (class 2; fig. 4c). At the level of the pixel grid, the frequency of whitings increases significantly with decreasing distance from the river mouth (n>40). The central and western parts of the lake show a very low frequency of whitings, < 10 occurrences counted between 2013 and 2021.

Figure 4d shows the temporal distribution of whitings (expressed in days of the year) for classes 1 and 2 observed between 2013 and 2021. Whitings of class 1 extend from late May to mid-September, with a peak in the second half of June. In contrast, the few whitings of class 2 occur in early August and early September (2 events of 3 and 4 days in August 2017 and

September 2021, respectively). The average AreaBGR values for both whiting types are shown in Figure 4e. The average AreaBGR index value for all events combined by class is about the same (~$1.6*10^4$). However, the whitings of class 1 show a higher range and outliers in the highest values (AreaBGR > $2.6*10^4$).

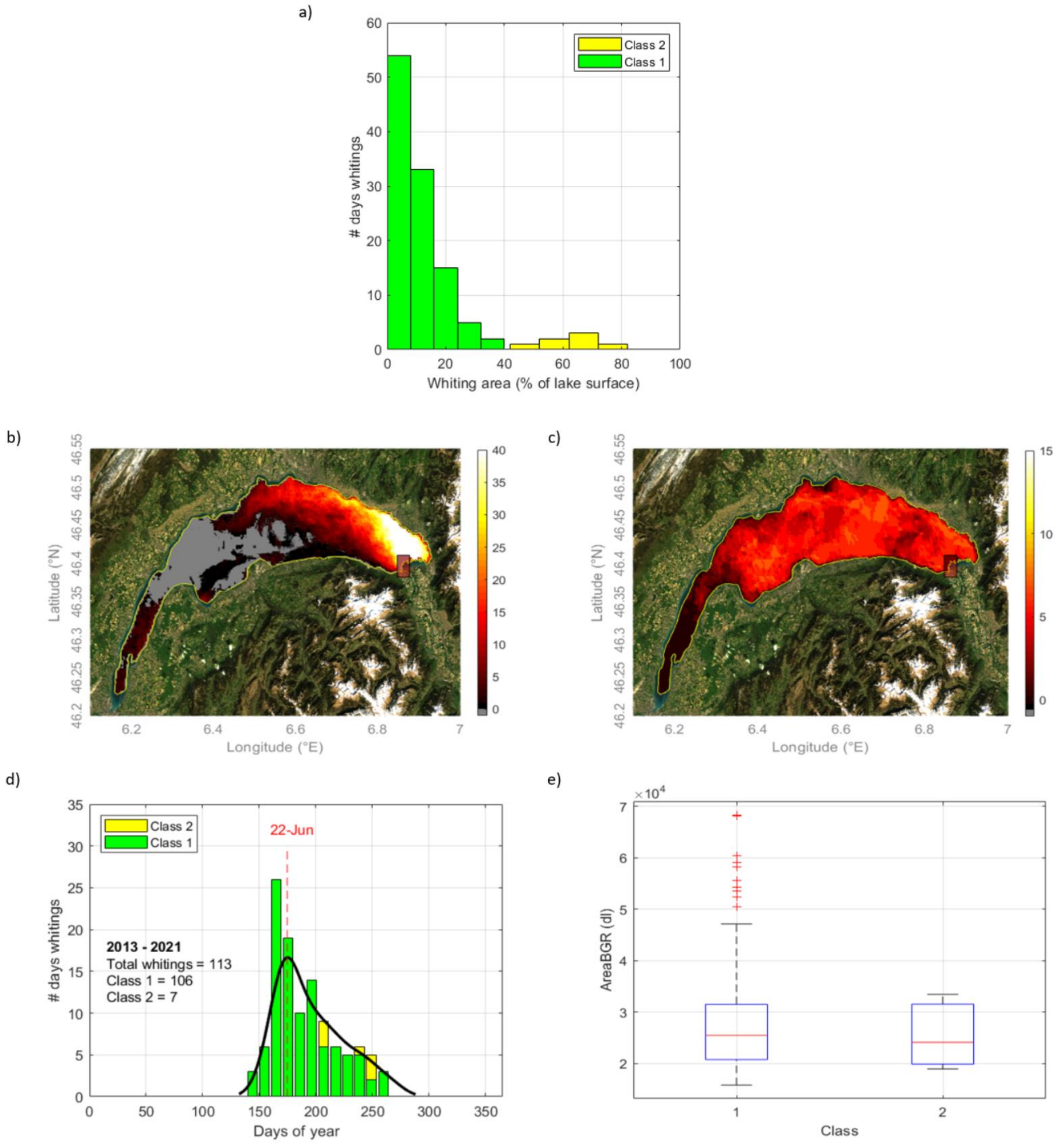


Figure 4. a) Distribution of whiting by areal coverage. Whitings with areas < 40% of the surface of the lake (class 1) and whitings with areas > 40 % of the lake (class 2). b-c) Maps of spatial occurrences of observed whiting days of class 1 (b) and the whitings of class 2 in Lake Geneva from 2013 to 2021 (c). d) Temporal distribution of whiting days (both class 1 (green) and class 2 (yellow) are stacked) expressed as a function of the day of the year of occurrence. The black line is the kernel fit. The date of the peak distribution is shown in red. e) Boxplots of the AreaBGR index values for whitings of classes 1 and 2. The median value is shown in red, and the top and

bottom edges of the box show the 25th and 75th percentiles. The black whiskers show extreme values, and the red crosses show outliers.

5.1.2 Temporal occurrences of observed whitings in Lake Geneva

The days of whiting and their spatial extent over 2013-2021, as detected from Landsat-8, Sentinel-2, and Sentinel-3 satellite images, are presented in Figure 5. Whitings are more frequently observed in 2018-2019 and 2021 (i.e., > 25 days) and reach greater maximal areas. Whiting days are less frequent in 2016 and 2017, and only three are detected in 2020. From 2013 to 2015, only the Landsat-8 dataset is available, the number of observations only represent a fraction of the later years, hence a much larger chance that whitings remain unseen (Fig. 5a).

Whitings of class 1 occur at high Rhone discharge (Fig. 5b, average discharge of about 320 $m^3\ s^{-1}$, Table 2) when air and water surface temperatures are high (i.e., approx. 22°C for air and 18°C between 0 and 10m for water, averaged over the observed whitings), and the thermocline depth is ca. 10m depth (Fig. 5c-e). Wind speed is more variable during the whiting days of class 1, with a mean value of 2.3 $m\ s^{-1}$ and a standard deviation of 1.2 $m\ s^{-1}$ (Fig. 5f). Whitings of class 2 occur in similar conditions, except for a lower Rhone discharge (i.e., approximately 250 $m^3\ s^{-1}$, Table 2). However, the limited number of class 2 events (i.e., only seven days) does not allow for further analysis.

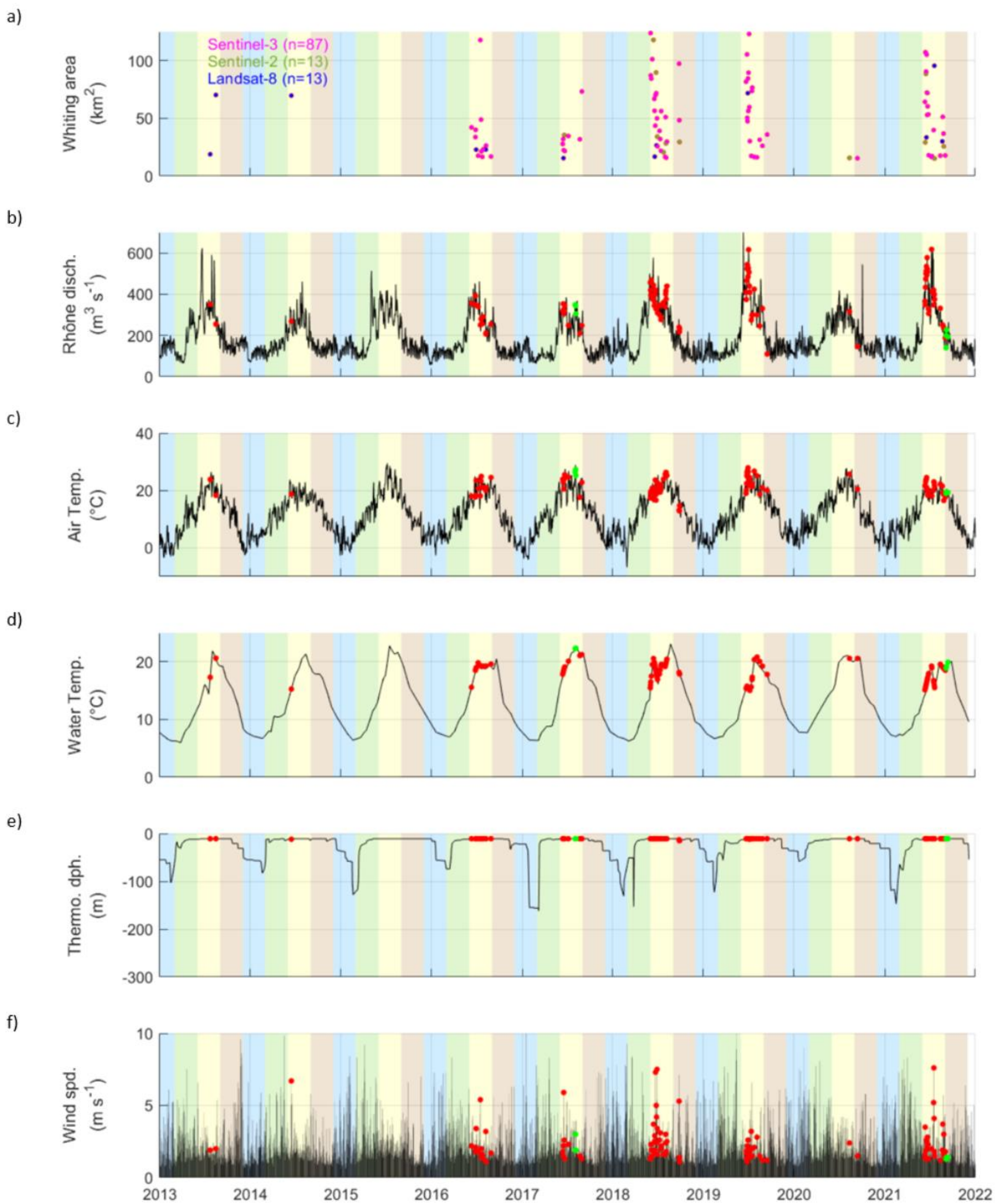


Figure 5. a) Whitings area as seen by the three satellites that detected events > 15 km^2. Landsat-8 images are shown in blue, Sentinel-2 in green, and Sentinel-3 in magenta. b) Rhone River discharge, c) air temperature (monitored at noon), d) surface water temperature (0-10m depth), e) thermocline depth at the SHL2 monitoring point, and f) wind speed. Shaded areas highlight the different seasons (blue for winter, green for spring, yellow for summer, and brown for fall). The occurrence of class 1 (class 2) whiting days is shown in red (green) on b-f.

Table 2. Averaged environmental conditions during observed whiting days from 2013 to 2021 (> 15 km^2) in Lake Geneva. The standard deviations for each condition are also specified. The number of whiting days for each class is specified.

	Class 1 (Mean +/- Std.)	Class 2 (Mean +/- Std.)
Rhone discharge ($m^3\ s^{-1}$)	363.1 +/- 102.9	251.1 +/- 81.4
Air temperature (°C)	21.5 +/- 3.0	22.3 +/- 3.8
Surface water temperature (°C)	17.9 +/- 1.8	20.6 +/- 1.7
Wind speed ($m\ s^{-1}$)	2.4 +/- 1.4	1.7 +/- 0.6
Thermocline depth (m)	11.1 +/- 0.6	11.0 +/- 0.0
# of obs. days	106	7

5.2 Machine learning and statistical approach

5.2.1 Drivers of whitings using machine learning

The detailed optimization results of the machine learning models are shown in Figure 6. The detailed method is described in the Supplementary Material. Note that only class 1 whiting days are considered, class 2 whitings being too few to be significantly related to the corresponding ground data.

As seen in section 2.6, the objective is to relate the occurrences of class 1 whitings to the corresponding ground data through the best model by comparing a DT and a RF algorithm. We first built a simple DT to determine the most important environmental factors to classify whiting events. The results show that water temperature and Rhone discharge are the two most discriminating factors for the occurrence of whitings between 2013 and 2021 (see Fig. 6a). Indeed, the two thresholds necessary to classify whitings are a minimum Rhone discharge of 207 $m^3\ s^{-1}$ and a minimum water temperature of 15°C. Using these thresholds allows for classifying the majority of the whitings (see the blue points in Figure 6b). This DT has good performances (validation AUC=0.86; validation accuracy = 74%), but can be improved by using the cost complexity pruning method. The best DT (see the Supplementary Material) has similar performances (validation AUC=0.83; validation accuracy = 81%), but still makes some classification errors by creating false positives (n=55 in the training dataset; n=28 in the validation dataset).

To go further, we compare the results obtained from the DT with those of the RF. The construction and optimization of the RF (see Supplementary Material) lead to the best RF composed of approximately twenty trees, with a training accuracy of ~1 (i.e., approx. 100% of whiting and non-whiting events in the training data have been correctly classified) and a validation AUC of 0.90. Besides, the model provides the most important indicators for the classification of whitings, namely Rhone discharge and water temperature (Fig. 6c). Using these two predictors and the decision boundaries, the classification results are shown in Figure 6d. The main advantage of this model is the consequent reduction of the number of false positives (n=0 in the training dataset; n=4 in the validation dataset) using a finer classification. This final RF is then used to reconstruct the past 'unseen' whiting days, based on the ground data monitored between 1958 and 2021 (see below).

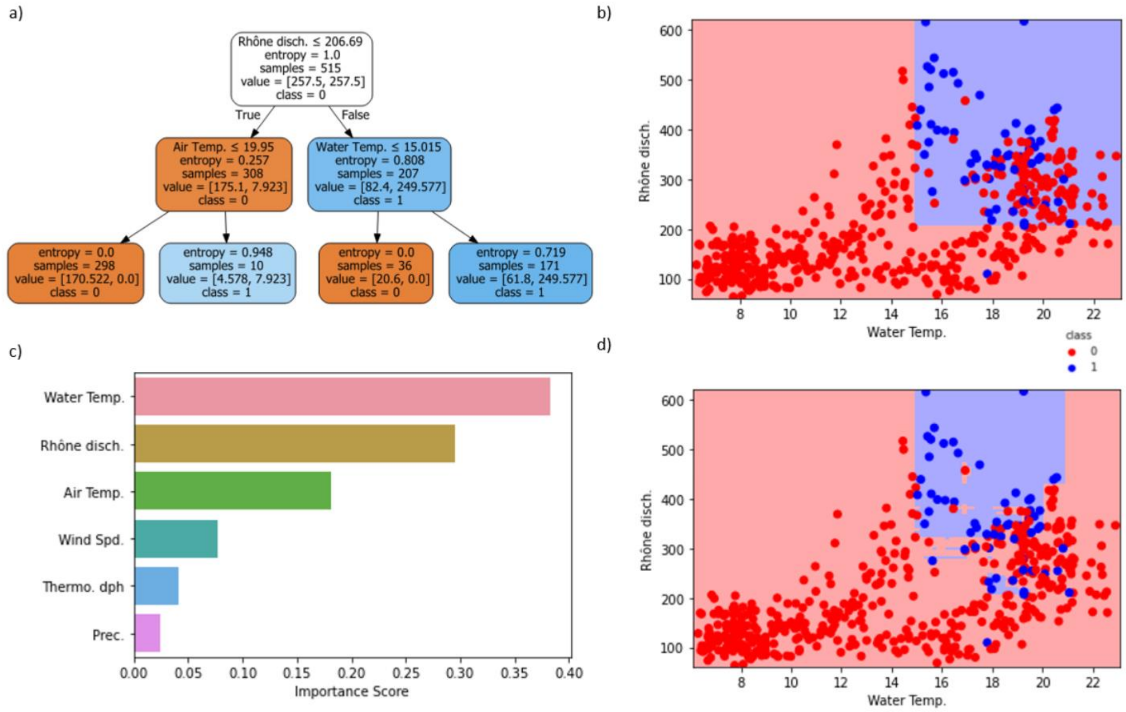


Figure 6. a) Decision tree used to classify the class 1 whiting days between 2013 and 2021 based on the surface water temperature (°C, 0-10m depth) and the Rhone discharge (m^3 s^{-1}). b) Results of the classification based on the DT shown in (a). Whitings and non-whitings events (from the training set) are shown as blue and red dots, respectively. The whitings and non-whitings decision zones predicted by the DT in (a) are the areas in blue and red, respectively. c) RF's important features. d) Same as (b) but using the RF algorithm with only the two most important features.

5.2.2 Reconstruction of past unseen whitings

Daily class 1 whiting presence-absence is reconstructed from the RF algorithm over the 1958-2021 time period (Fig. 7a). The total number of whitings (class 1, expressed as days per year) is highly variable over the years (annual average of n = 18 days of whiting per year). Values range from years with very few or no whiting days (n<3; 1964, 1974, 1976, 1997) to years with frequent whiting days (n > 35; 1958, 1963, 1966, 1982, 1994, 2001) (Fig. 7a). Neither the Mann-Kendall test ($p_{M\text{-}K}$=0.117) nor the BEAST decomposition (low probability of changing points) detect any clear temporal trend in the annual whiting occurrence between 1958 and 2021, reconstructed by the RF algorithm (Fig. 7b). There is yet a shift in the whiting phenology. The number of spring whiting increases from 1958 to 2021 ($p_{M\text{-}K}$=0.011; Fig. 7c). The BEAST decomposition detected a changing point in 2000 (maximum probability in changing points). It corresponds to an increase in spring whiting occurrence (+1 day on average since 2000).

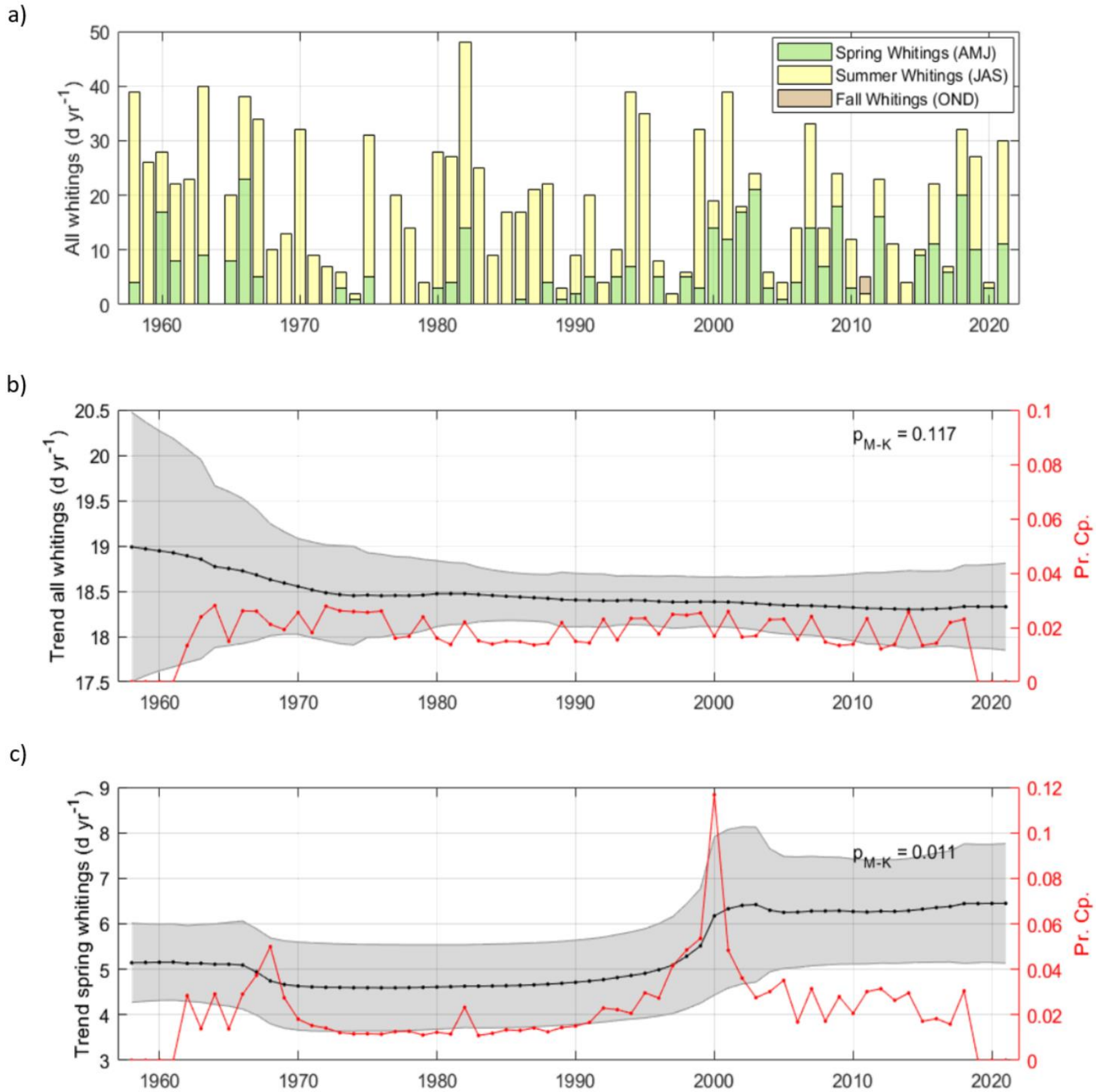


Figure 7. a) Reconstruction of past 'unseen' whiting days from 1958 to 2021 (Class 1 only, > 15 km^2, expressed as days of whiting per year) based on the RF algorithm. Whitings are separated following the season of occurrence (green: spring, yellow: summer, brown: fall). b) BEAST decomposition of the time series of all whitings, with the trend indicated in black and the standard deviation in grey. The probability of finding a changing point in the trend is shown in red. c) same as b) but for spring whiting days only. The p-values of the Mann-Kendall tests are specified on b) and c).

5.2.3 Factors controlling occurrences of whitings from 1958 to 2021

Here we attempt to determine the relationship between the temporal variability of class 1 whiting occurrences in Lake Geneva and climatic indices. The interannual and seasonal variabilities of whiting days reconstructed from the RF algorithm are tested against the climate indices that affect central Europe and Switzerland.

The inter-annual variability of the total and spring numbers of whitings (expressed as anomalies in days per year) is shown in Figure 8. A comparison is made between the whiting anomalies per year, using the RF algorithm predictions, and the climatic indices most known to influence the Swiss and European climates. The anomalies in the total number of whiting days per year can be partly explained by the climatic index WeMOI (Fig. 8a; r=0.36, p-value of 0.004). Besides, the anomalies of spring whiting days are related to the AMO index (Fig. 8b; r=0.33, p-value of 0.007). The other climate indices (such as NAO and ONI) do not seem to correlate significantly with the interannual changes in the total, nor spring, numbers of whiting days. Over the period represented (1958-2021), positive WeMOI values tend to increase the total number of whiting days per year. On the contrary, negative WeMOI values tend to reduce the total number of whiting days per year. A similar observation can be made with the AMO index, which tends to increase the number of spring whitings while positive. The period when spring whitings are minimal (between 1967 and 2000) corresponds mainly to negative AMO values.

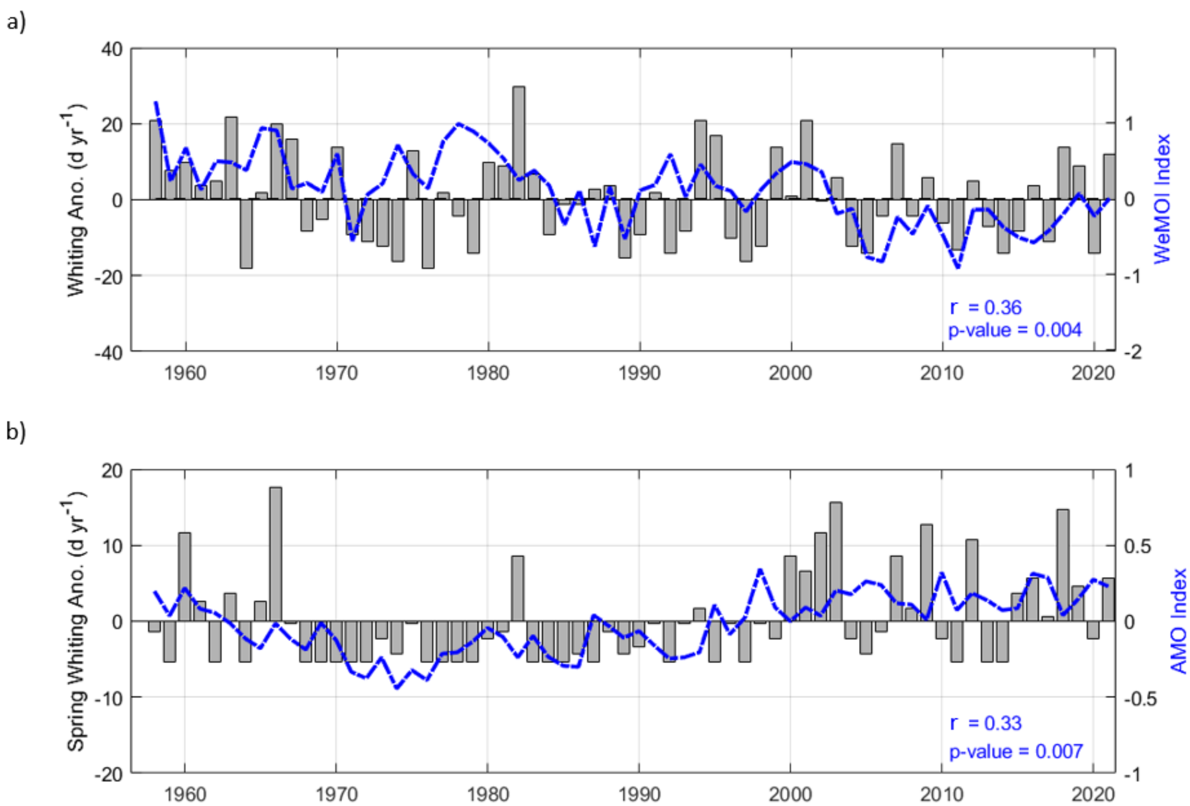


Figure 8. a) Time series of annual predicted whiting events anomalies (class 1) using the RF algorithm and correlated Western Mediterranean Oscillation index (WeMOi, r=0.36). b) same as a) but for spring whiting days only, correlated with the Atlantic Multidecadal Oscillation (AMO, r=0.33).

6 Discussion and conclusions

The objective of this study is to measure the spatial extent and temporal occurrences of whiting days (i.e., massive clouds of suspended $CaCO_3$ particles induced by intense calcite precipitation) in Lake Geneva using Landsat-8, Sentinel-2, and Sentinel-3 satellite data between

2013 and 2021. A RF algorithm then demonstrates the link between these occurrences and the meteorological, lake physical, and riverine conditions. The latter is finally used to reconstruct the past occurrences between 1958 and 2021 based on the main identified controlling factors of whitings, namely the Rhone River discharge and the lake surface water temperature. Below we first discuss the complementarity of the satellites and the robustness of the index used. Then, we detail the results obtained regarding spatial and temporal observations and discuss the reconstruction of past whiting days in light of the climatic indices influencing the central part of Europe.

6.1 Remote sensing of whitings in Lake Geneva

Satellite observations are increasingly used to characterize biogeochemical processes in inland waters (Verpoorter et al., 2014; Spyrakos et al., 2020; Seegers et al., 2021). We chose to combine Sentinel-2 and Landsat-8 datasets with Sentinel-3 to describe whitings in Lake Geneva. The different spatial (i.e., 30m or 300m) and temporal (i.e., 1 day or approx. 15 days) resolutions enable a relatively good monitoring of the aspect of Lake Geneva over the period 2013-2021.We observe different responses on the Landsat-8, Sentinel-2, and Sentinel-3 data due to various product sources and processes. The inter-calibration carried out in this work expresses the satellite responses in term of AreaBGR in the same range, which is needed for the time series coherence (Fig. 2).

We use the AreaBGR index to detect whiting days in Lake Geneva. Indeed, intense events of $CaCO_3$ precipitation lead to an increase in the water reflectance, mainly in the green band, resulting in a turquoise watercolor. This result contrasts sharply with the lake's color without precipitation, which appears dark in the visible spectrum (Heine et al., 2017). This index responds positively to various suspended particles (sediments and phytoplankton species) that influence the visible spectrum by backscattering sunlight (see 4.1). Among these suspended particles, distinguishing the sedimentary contributions from the Rhône (i.e., inputs that reach the surface when the lake is unstratified) and resuspension by near-shore waves, from the precipitation of $CaCO_3$ particles during whitings can be challenging. The use of specific filters, determined from geochemical knowledge about the whiting process, enables building a conservative database retaining only whiting days. Although empirical, these filters could be further tested on different peri-alpine lakes to build a process chain for validating the AreaBGR index as a proxy of whitings.

Besides, we do not use specific filters related to the presence of phytoplankton in the lake. Indeed, some biological blooms can potentially influence the reflectance used to calculate AreaBGR, without inducing whiting events. However, their abundance in Lake Geneva is never high enough to reach the AreaBGR threshold and we did not find an example of this contamination in our database in line with the study of Nouchi et al. (2019). The ongoing development of remote sensing monitoring of primary production and phytoplankton species is crucial to better characterize the possible contamination of the AreaBGR index from organic sources.

6.2 Spatial and temporal occurrences of whitings in Lake Geneva

The majority of whitings in Lake Geneva tends to occur during early summer while fewer events occurred later during the season (Fig. 4). These two types are associated to different spatial patterns. Thus, the determinism of these two classes can be related and explained by diverse environmental drivers, notably identified through machine learning techniques for the majority of them (class 1 whitings), and are probably triggered by different mechanisms of nucleation. Indeed, the spatial extent of the majority of whiting days tends to be related to the Rhône inflow (>95 %, see Fig. 4b).The turbidity inputs of the Rhône can trigger the nucleation of $CaCO_3$ particles during high discharge when the lake is stratified, and the surface water temperature is high. This result is in line with the previous works of Nouchi et al. (2019) and Escoffier et al. (2022). Authors highlighted the role of the interflow in triggering whiting events when the spread of fine sediments along the whole lake is driven by local hydrodynamics during the high physical stability of the water column (Giovanoli, 1990; Cotte and Vennemann, 2020). Detrital $CaCO_3$ particles eroded from the watershed could also participate in whitings detection close to the River mouth (Escoffier et al., 2022), increasing the reflectance of surface waters and the AreaBGR mean and extreme values (see Fig. 4e).

However, fewer class 2 whiting events are detected in the central part of the lake (i.e. approx. 5% in the period 2013-2021), later during the season. The lack of in situ measurements during those whitings and the few events observed do not allow a more refined characterization. They can probably be related to episodes of important primary production, i.e. phytoplankton bloom in early August 2017 (CIPEL Report, 2018), and a massive, transient *Uroglena sp.* bloom in Sept 2021 (UMR CARRTEL INRAE USMB, 2021). The influence of primary production in triggering whiting events is still under debate and can be considered in several ways. Primary production tends to increase pH and favor calcite supersaturation and potential precipitation. However, the nucleation of calcite particles during precipitation can occur on small picoplankton cells (Dittrich and Obst, 2004) but also on algal-derived exopolymeric substances (EPS) or other suitable heteronuclei (bacteria). Moreover, as discussed before, high levels of chlorophyll a during phytoplankton blooms can also influence the AreaBGR index and potentially bias the corresponding whiting detection. Coupling in situ measurements of primary production and characterization of phytoplankton species in line with $CaCO_3$ measurements could provide crucial information on the biologically induced precipitation of calcite. A future study should also compare a lake under the influence of a glacial river, i.e. subject to turbid inputs (such as Lake Geneva), to a lake without glacial inputs but where whiting events are observed (Lake Neuchâtel). The study of the difference in spatial and temporal occurrences could reveal different roles of organic and inorganic processes in the triggering of whiting events.

6.3 The long-term evolution of whitings in Lake Geneva

We reconstruct the class 1 whiting occurrences, as days per year, between 1958 and 2021, based on the RF algorithm (Fig. 7). The number of reconstructed whiting days per year is very variable, with no noticeable trend in its long-term evolution. However, the interannual variability can be partly related to the WeMOi (Fig. 8a). This index is causally related to precipitation in northern Italy, which could be at the origin of environmental conditions in Switzerland, especially in precipitation changes over years that could impact Rhone River discharge and related turbid inputs to Lake Geneva. Mediterranean climatic activity thus seems to play a role in changes in the total number of whiting events per year. When the WeMOi is high, whiting days related to Rhone River inputs (i.e. the 95% of total events in our case) are more frequent.

In addition, we observe a seasonal trend with the increase of early whitings since 2000 (Fig. 7c). This change coincides with a change in climate regime due to the AMO (Fig. 8b). Indeed, the positive values of the index since 2000 and the observed upward trend show the general increase in temperatures measured in Europe (Knight et al., 2006). The latter changes the Swiss climate, and the physical conditions of the lake, especially the temperature and stratification of the surface water that warmed and stratified earlier in the year. The conditions necessary for the onset of whitings in Lake Geneva are therefore met earlier in the year, in terms of Rhone River inputs, water temperature, and water column stratification.

Although our study significantly quantified the inter-annual variability in the total number of whiting events and the trend in their phenology (p-values <0.01), correlation coefficients of only 0.36 and 0.33 respectively have been obtained (Fig. 8). Other environmental, region-specific factors probably actively participate in the inter-annual change in whiting occurrences. Among them, the increase in alkalinity and Ca2+ concentration of the Rhône over last decades (Zobrist et al., 2018), as well as changes in discharge and sediment load related to human activities (Lane et al., 2019) could be at the origin of an additional variability that cannot be quantified from climatic indices.

To go further, future changes in Mediterranean and Atlantic activities related to global warming could influence environmental conditions in Switzerland. The trend in the number of whiting days per year depends on the Rhône discharge, impacted mainly by precipitation, snow- and ice melt. Based on the work of Freudiger et al. (2020), the annual Rhône discharge could remain stable in the future (2020-2100), leading to a total number of whitings that does not follow a specific trend, but from whose annual changes are in line with the WeMOi. However, the contribution of the Rhône discharge could highly change with an increase in rainfall, related to a decrease in the snow- and ice melt induced by earlier warmer temperatures. This could cause a change in the peak discharge of the Rhône with maximal discharges met earlier in the year. On the other hand, higher water temperatures may positively act on calcite supersaturation (due to its retrograde solubility). The periods of calcite supersaturation and lake stratification may start earlier and last longer. All this may change the relative influence of the environmental drivers identified in this work, with a change in whiting phenology and abundances of class 1 vs class 2 whitings in Lake Geneva, in line with changes in AMO.

This shift in whiting phenology could have several consequences on the functioning of the lake ecosystem. First, as whitings increase lake surface turbidity, light-dependent processes such as spring phytoplankton blooms could be altered. Earlier whitings could decrease the intensity of light received during these crucial bloom periods (Long et al. 2017; 2018). In addition, the carbon transfer to the benthic layer in the form of calcite actively participates in nutrient cycling. It appears crucial to estimate the impact that climate change may have on the future evolution of the frequency of whitings. The role of these events in the annual $CaCO_3$ precipitation and its transfer to the benthic ecosystem and the burial of carbon remains to be determined.

Acknowledgments

This study was supported by the CARBOGEN project (SNF 200021_175530).

References

Binding, C. E., Greenberg, T. A., Watson, S. B., Rastin, S., & Gould, J. (2015). Long term water clarity changes in North America's Great Lakes from multi-sensor satellite observations. *Limnology and Oceanography*, *60*(6), 1976-1995.

Copernicus. (2020). Copernicus global land operations "cryosphere and water" "CGLOPS-2"—Algorithm theoretical basis document lake waters 300m and 1km products versions 1.3.0-1.4.0 Issue I1.12.

Cotte, G., & Vennemann, T. W. (2020). Mixing of Rhône River water in Lake Geneva: Seasonal tracing using stable isotope composition of water. *Journal of Great Lakes Research*, *46*(4), 839-849.

Dierssen, H. M., Zimmerman, R. C., & Burdige, D. J. (2009). Optics and remote sensing of Bahamian carbonate sediment whitings and potential relationship to wind-driven Langmuir circulation. *Biogeosciences*, *6*(3), 487-500.

Dittrich, M., & Obst, M. (2004). Are picoplankton responsible for calcite precipitation in lakes?. *AMBIO: A Journal of the Human Environment*, *33*(8), 559-564.

Dittrich, M., Kurz, P., & Wehrli, B. (2004). The role of autotrophic picocyanobacteria in calcite precipitation in an oligotrophic lake. *Geomicrobiology Journal*, *21*(1), 45-53.

Effler, S. W. (1987). The importance of whiting as a component of raw water turbidity. *Journal-American Water Works Association*, *79*(3), 80-82.

Effler, S. W., & Peng, F. (2012). Light-scattering components and Secchi depth implications in Onondaga Lake, New York, USA. *Fundamental and applied limnology*, *179*(4), 251-265.

ESA. (2020). D2.2: Algorithm theoretical basis document.

Escoffier, N., Perolo, P., Lambert, T., Rüegg, J., Odermatt, D., Adatte, T., ... & Perga, M. E. (2022). Whiting events in a large peri-alpine lake: Evidence of a catchment-scale process. *Journal of Geophysical Research: Biogeosciences*, e2022JG006823.

FOEN. (2022). Commander des données hydrologiques historiques et validées. https://www.bafu.admin.ch/bafu/fr/home/themes/eaux/etat/donnees/obtenir-des-donnees-mesurees-sur-le-theme-de-l-eau/commander-des-donnees-hydrologiques-historiques-et-validees.html

Freudiger, D., Vis, M., and Seibert, J. (2020). Quantifying the contributions to discharge of snow and glacier melt. Hydro-CH2018 project. Commissioned by the Federal Office for the Environment (FOEN), Bern, Switzerland, 49 pp.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. " O'Reilly Media, Inc.".

Giovanoli, F. (1990). Horizontal transport and sedimentation by interflows and turbidity currents in Lake Geneva. In *Large lakes* (pp. 175-195). Springer, Berlin, Heidelberg.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, *135*(2-3), 147-186.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (Vol. 2, pp. 1-758). New York: springer.

Heine, I., Brauer, A., Heim, B., Itzerott, S., Kasprzak, P., Kienel, U., & Kleinschmit, B. (2017). Monitoring of calcite precipitation in hardwater lakes with multi-spectral remote sensing archives. *Water*, *9*(1), 15.

Hodell, D. A., Schelske, C. L., Fahnenstiel, G. L., & Robbins, L. L. (1998). Biologically induced calcite and its isotopic composition in Lake Ontario. *Limnology and Oceanography*, *43*(2), 187-199.

Izquierdo Miguel, R., Alarcón Jordán, M., & Avila Castells, A. (2014). WeMO effects on the amount and the chemistry of winter precipitation in the north-eastern Iberian Peninsula. *Tethys: Journal of Mediterranean Meteorology and Climatology*, *10*, 45-51.

Kendall, M. G. (1948). Rank correlation methods.

Khan, H., Marcé, R., Laas, A., & Obrador Sala, B. (2022). The relevance of pelagic calcification in the global carbon budget of lakes and reservoirs. *Limnetica, 2022, vol. 41*.

Knight, J. R., Folland, C. K., & Scaife, A. A. (2006). Climate impacts of the Atlantic multidecadal oscillation. *Geophysical Research Letters*, *33*(17).

Kumar, L., & Mutanga, O. (2018). Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sensing*, *10*(10), 1509.

Lambert, T., & Perga, M. E. (2019). Non-conservative patterns of dissolved organic matter degradation when and where lake water mixes. *Aquatic Sciences, 81*(4), 1-13.

Lane, S. N., Bakker, M., Costa, A., Girardclos, S., Loizeau, J. L., Molnar, P., ... & Schlunegger, F. (2019). Making stratigraphy in the Anthropocene: climate change impacts and economic conditions controlling the supply of sediment to Lake Geneva. *Scientific reports*, *9*(1), 1-11.

Larson, E. B., & Mylroie, J. E. (2014). A review of whiting formation in the Bahamas and new models. *Carbonates and Evaporites, 29*(4), 337-347.

Lloyd-Hughes, B., & Saunders, M. A. (2002). Seasonal prediction of European spring precipitation from El Niño–Southern Oscillation and local sea-surface temperatures. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *22*(1), 1-14.

Loizeau, J. L., & Dominik, J. (2000). Evolution of the Upper Rhone River discharge and suspended sediment load during the last 80 years and some implications for Lake Geneva. *Aquatic Sciences*, *62*(1), 54-67.

Long, J. S., Hu, C., Robbins, L. L., Byrne, R. H., Paul, J. H., & Wolny, J. L. (2017). Optical and biochemical properties of a southwest Florida whiting event. *Estuarine, Coastal and Shelf Science*, *196*, 258-268.

Long, J. S., Hu, C., & Wang, M. (2018). Long-term spatiotemporal variability of southwest Florida whiting events from MODIS observations. *International Journal of Remote Sensing*, *39*(3), 906-923.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the econometric society*, 245-259.

Martin-Vide, J., & Lopez-Bustins, J. A. (2006). The western Mediterranean oscillation and rainfall in the Iberian Peninsula. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *26*(11), 1455-1475.

McPhaden, M. J., Zebiak, S. E., & Glantz, M. H. (2006). ENSO as an integrating concept in earth science. *science*, *314*(5806), 1740-1745.

Molinero, J. C., Anneville, O., Souissi, S., Lainé, L., & Gerdeaux, D. (2007). Decadal changes in water temperature and ecological time series in Lake Geneva, Europe—relationship to subtropical Atlantic climate variability. *Climate Research*, *34*(1), 15-23.

Müller, B., Meyer, J. S., & Gächter, R. (2016). Alkalinity regulation in calcium carbonate-buffered lakes. *Limnology and Oceanography*, *61*(1), 341-352.

Muller-Wilm, U., Louis, J., Richter, R., Gascon, F., & Niezette, M. (2013). Sentinel-2 level 2A prototype processor: Architecture, algorithms and first results. In *Proceedings of the ESA Living Planet Symposium, Edinburgh, UK* (pp. 9-13).

Mutanga, O., & Kumar, L. (2019). Google earth engine applications. *Remote Sensing*, *11*(5), 591.

Nouchi, V., Kutser, T., Wüest, A., Müller, B., Odermatt, D., Baracchini, T., & Bouffard, D. (2019). Resolving biogeochemical processes in lakes using remote sensing. *Aquatic Sciences*, *81*(2), 27.

Peng, F., & Effler, S. W. (2011). Characterizations of the light-scattering attributes of mineral particles in Lake Ontario and the effects of whiting. *Journal of Great Lakes Research*, *37*(4), 672-682.

Peng, F., & Effler, S. W. (2017). Characterizations of calcite particles and evaluations of their light scattering effects in lacustrine systems. *Limnology and Oceanography*, *62*(2), 645-664.

Perga, M. E., Maberly, S. C., Jenny, J. P., Alric, B., Pignol, C., & Naffrechoux, E. (2016). A century of human-driven changes in the carbon dioxide concentration of lakes. *Global Biogeochemical Cycles*, *30*(2), 93-104.

Ridgwell, A., & Zeebe, R. E. (2005). The role of the global carbonate cycle in the regulation and evolution of the Earth system. *Earth and Planetary Science Letters*, *234*(3-4), 299-315.

Rimet, F., Anneville, O., Barbet, D., Chardon, C., Crepin, L., Domaizon, I., ... & Monet, G. (2020). The Observatory on LAkes (OLA) database: Sixty years of environmental data accessible to the public. *Journal of Limnology*, *79*(2), 164-178.

Schwefel, R., Gaudard, A., Wüest, A., & Bouffard, D. (2016). Effects of climate change on deepwater oxygen and winter mixing in a deep lake (Lake Geneva): comparing observational findings and modeling. *Water Resources Research*, *52*(11), 8811-8826. https://doi.org/10.1002/2016WR019194

Seegers, B. N., Werdell, P. J., Vandermeulen, R. A., Salls, W., Stumpf, R. P., Schaeffer, B. A., ... & Loftin, K. A. (2021). Satellites for long-term monitoring of inland US lakes: The MERIS time series and application for chlorophyll-a. *Remote Sensing of Environment*, *266*, 112685.

Shanableh, A., Al-Ruzouq, R., Gibril, M. B. A., Flesia, C., & Al-Mansoori, S. (2019). Spatiotemporal mapping and monitoring of whiting in the semi-enclosed gulf using moderate resolution imaging spectroradiometer (MODIS) time series images and a generic ensemble tree-based model. *Remote Sensing*, *11*(10), 1193.

Shanableh, A., Al-Ruzouq, R., Gibril, M. B. A., Khalil, M. A., AL-Mansoori, S., Yilmaz, A. G., ... & Flesia, C. (2021). Potential Factors That Trigger the Suspension of Calcium Carbonate Sediments and Whiting in a Semi-enclosed Gulf. *Remote Sensing*, *13*(23), 4795.

Sondi, I., & Juracic, M. (2010). Whiting events and the formation of aragonite in Mediterranean Karstic Marine Lakes: new evidence on its biologically induced inorganic origin. *Sedimentology*, *57*(1),85-95.

Stabel, H. H. (1986). Calcite precipitation in Lake Constance: Chemical equilibrium, sedimentation, and nucleation by algae 1. *Limnology and Oceanography*, *31*(5), 1081-1094.

Steinmetz, F., & Ramon, D. (2018). Sentinel-2 MSI and Sentinel-3 OLCI consistent ocean colour products using POLYMER. In *Remote sensing of the open and coastal ocean and inland waters* (Vol. 10778, pp. 46-55). SPIE.

Strong, A. E., & Eadie, B. J. (1978). Satellite observations of calcium carbonate precipitations in the Great Lakes 1. *Limnology and Oceanography*, *23*(5), 877-887.

Spyrakos, E., Hunter, P., Simis, S., Neil, C., Riddick, C., Wang, S., ... & Tyler, A. (2020). Moving towards global satellite based products for monitoring of inland and coastal waters. Regional examples from Europe and South America. In *2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)* (pp. 363-368). IEEE.

Thompson, J. B., Schultze-Lam, S., Beveridge, T. J., & Des Marais, D. J. (1997). Whiting events: biogenic origin due to the photosynthetic activity of cyanobacterial picoplankton. *Limnology and oceanography*, *42*(1), 133-141.

UMR CARRTEL INRAE USMB. (2021). Bienvenue à l'UMR CARRTEL. Bloom d'une microalgue, sur le Léman : Uroglena sp. non toxique, qui colore l'eau en marron. inra.com. https://www6.lyon-grenoble.inrae.fr/carrtel_fre/Centre-Alpin-de-Recherche-sur-les-Reseaux-Trophiques-des-Ecosystemes-Limniques/Actualites-CARRTEL-INRAE-USMB/Bloom-d-une-microalgue-sur-le-Leman-Uroglena-sp.-non-toxique-qui-colore-l-eau-en-marron

Verpoorter, C., Kutser, T., Seekell, D. A., & Tranvik, L. J. (2014). A global inventory of lakes based on high-resolution satellite imagery. *Geophysical Research Letters*, *41*(18), 6396-6402.

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., ... & Zhu, Z. (2019). Current status of Landsat program, science, and applications. *Remote sensing of environment*, *225*, 127-147.

Zhao, K., Wulder, M. A., Hu, T., Bright, R., Wu, Q., Qin, H., ... & Brown, M. (2019). Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote sensing of Environment*, *232*, 111181.

Zobrist, J., Schoenenberger, U., Figura, S., & Hug, S. J. (2018). Long-term trends in Swiss rivers sampled continuously over 39 years reflect changes in geochemical processes and pollution. *Environmental Science and Pollution Research*, *25*(17), 16788-16809.