A diatom-based palaeoprecipitation record for Bonaire

Romée van der Kuil[2], Kees Nooren[1,2], Francesca Sangiorgi[2], Madlene Nussbaum[2], Timme Donders[2]
[1]Royal Netherlands Institute of Southeast Asian and Caribbean Studies (KITLV/KNAW), 2311 BE Leiden
[2]Faculty of Geosciences, Utrecht University, 3584 CB Utrecht

Introduction

This research aims to create a better understanding of the frequency and duration of past droughts and their impact on Bonaire. For this purpose I will use diatoms as a proxy for drought, as they are very sensitive to salinity changes and therefore a powerful tool for reconstructing a palaeoprecipitation record of the late holocene. To create a high-resolution record (>300 samples) manual counting is extremely time-consuming. Therefore I will develop a method to automatically count diatoms from virtual slides using recent advancements in deep learning.

Area and Core

Saliña Bartol is a hypersaline lagoon, separated from the sea by a coral rubble ridge. We constructed a raft using pallets and tractor tires to reach the middle of the lagoon, and took two cores using a piston corer.

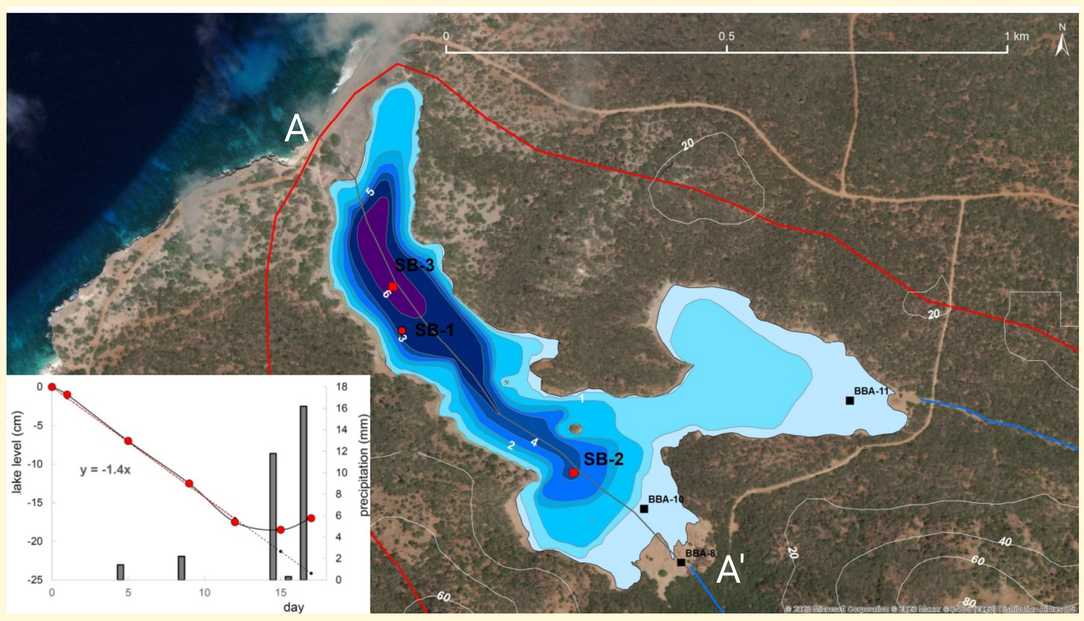

Figure 1. Map of Saliña Bartol including the locations of the gravity core (SB-1), cores taken with the piston corer (SB-2, SB-3), and lake level/rainfall between November 23 and December 10, 2023.

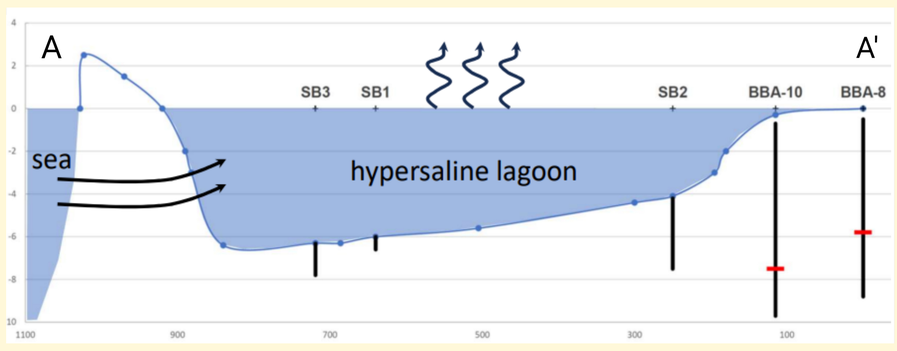

Figure 2. Transect A- A'. Our cores (SB1-SB3) were not deep enough to encounter the tsunami deposit (red bar in core BBA-8 and BBA-10) found by Engel et al. (2013)[3]. The tsunami deposit was dated to approximately 3000 yr BP.

The core that is used for this research is SB-2, which is the longest, and likely spans the past 2000 years.

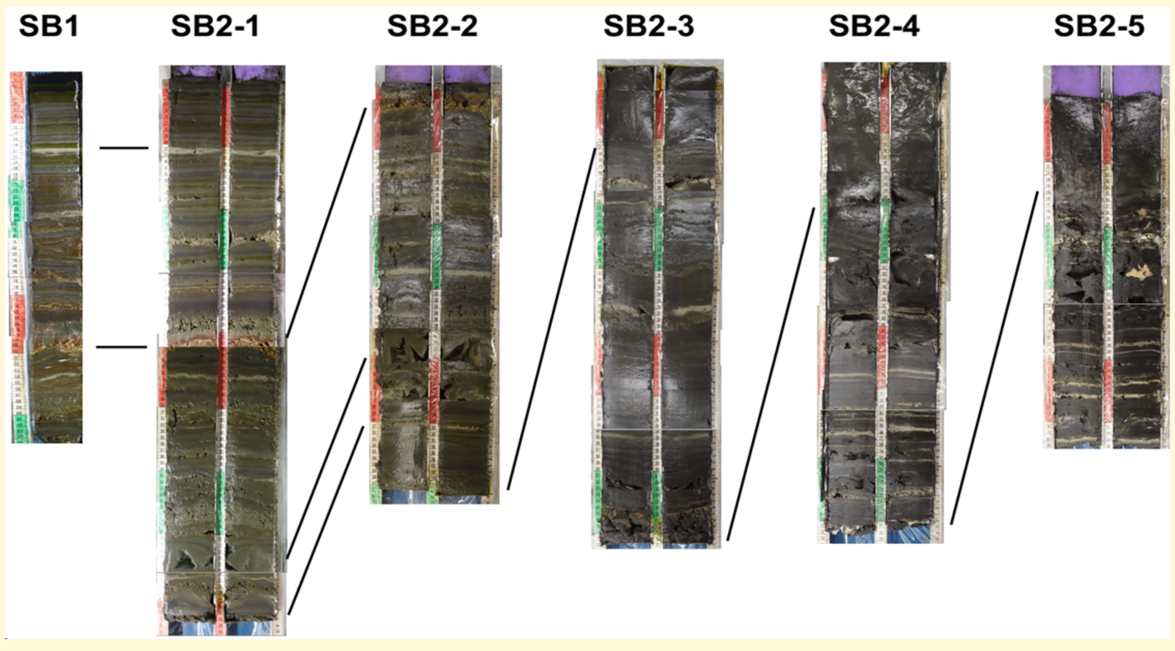

Figure 3. The gravity core (SB1) and one of the cores taken with the piston corer (SB2-1, SB2-2, SB2-3, SB2-4, SB2-5).

Deep Learning for Diatom Classification

A slide-scanning microscope is used to create virtual diatom slides. To train the model a training dataset will be created by annotating diatom slides in "Pollenlabeler"[4] (fig. 4). To get the coordinates from the diatoms in the pictures edge-detection software SHERPA is used[5]. This data is then used to train the Yolov8 image segmentation and classification model developed by Ultralytics[6].

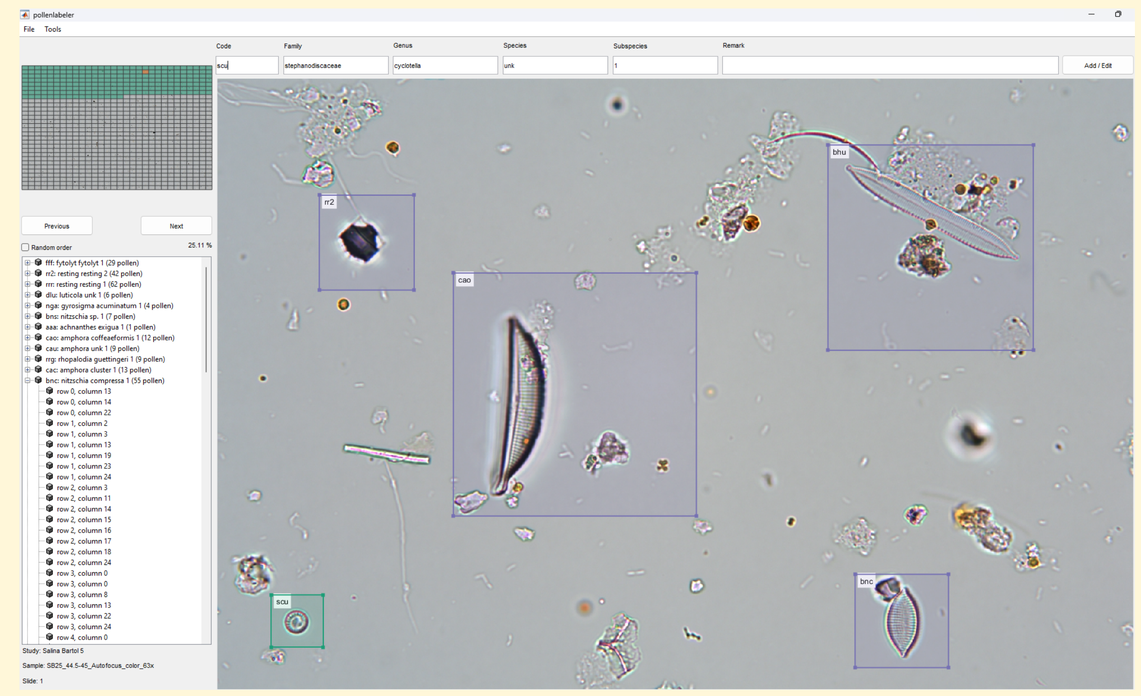

Figure 4. Annotation of diatoms using the pollenlabeler.

A problem with using deep learning is that models are usually trained on small image sizes (640x640 pixels), while our virtual slides are very big. To deal with this I will use SAHI[7] (Slicing Aided Hyper Inference), which slices the large pictures into tiles to draw inference, and stitches the results back together.

Figure 5. Segmentation of small objects (birds) in a high resolution image without SAHI and with SAHI (source: learnopencv.com).

Preliminary Findings

So far 14 of the 80 diatom slides from SB2-5 have been counted manually. Numerous diatom species were found of which a few tend to be more dominant. $CaCO_3$ content and LOI (Loss on Ignition) of SB2-5 and SB2-4 were also measured.

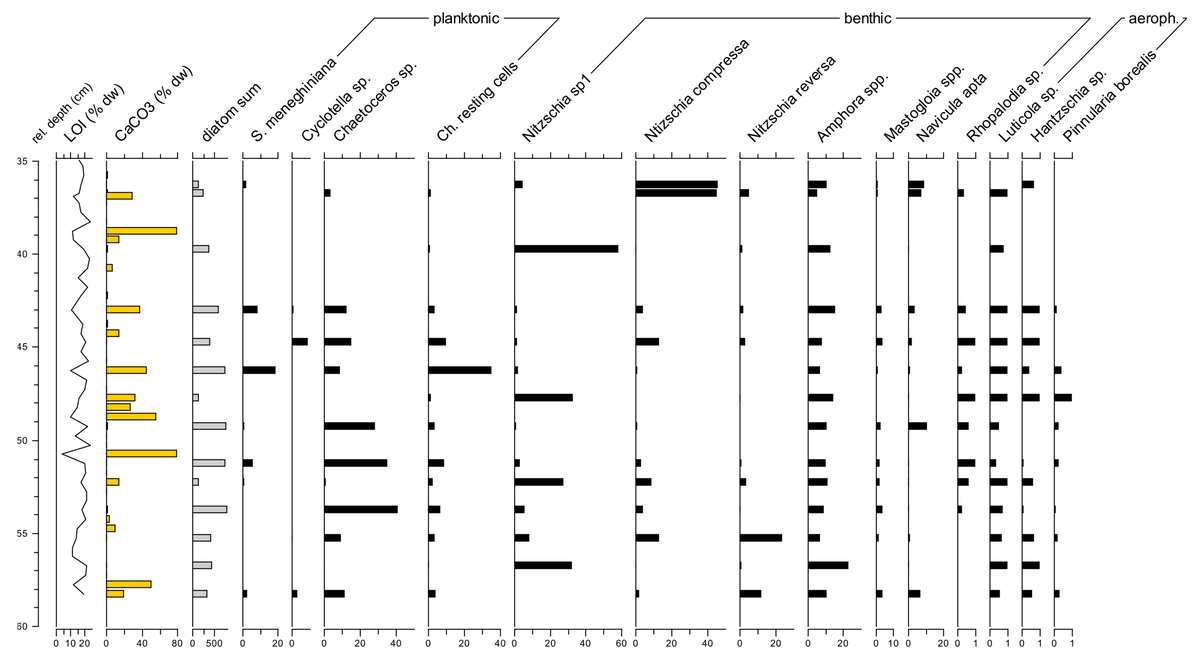

Figure 6. LOI, $CaCO_3$ content, and diatom relative abundance (%) for the most common species. The diatom species in these manually counted samples are used as a training set for deep learning.

[3]Engel, M., Brückner, H., Fürstenberg, S., Frenzel, P., Konopczak, A.M., Scheffers, A., Kelletat, D., May, S.M., Schäbitz, F., Daut, G. (2013). A prehistoric tsunami induced long-lasting ecosystem changes on a semi-arid tropical island – the case of Boka Bartol (Bonaire, Leeward Antilles). Naturwissenschaften 100, 51–67.
[4]Developed by E. Bennink UU/UMC
[5]Kloster, M., Kauer, G. & Beszteri, B. SHERPA: an image segmentation and outline feature extraction tool for diatoms and other objects. *BMC Bioinformatics* **15**, 218 (2014).
[6]Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLO (Version 8.0.0) [Computer software]. https://github.com/ultralytics/ultralytics
[7]Akyon, F. C., Altinuc, S. O., & Temizel, A. (2022). Slicing Aided Hyper Inference and Fine-tuning for Small Object Detection. 2022 IEEE International Conference on Image Processing (ICIP), 966–970. https://doi.org/10.1109/ICIP46576.2022.9897990