Start Clim 2021

Fernerkundungsbasiertes Monitoring und datengetriebene Modellierung der Wasserflächen im Nationalpark Neusiedler See – Seewinkel (FEMOWINKEL)

Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie

Bundesministerium Bildung, Wissenschaft und Forschung

klima+ energie fonds

LAND OBERÖSTERREICH

ober österreich

umweltbundesamt

BOKU

Autoren: Henri Schauer, B.Sc.; Dr. Stefan Schlaffer; Emanuel Büechi, M.Sc.; Prof. Dr. Wouter Dorigo

Diese Publikation sollte folgendermaßen zitiert werden:

Schauer, H.; Schlaffer, S.; Büechi, E.; Dorigo, W. (2022): Fernerkundungsbasiertes Monitoring und datengetriebene Modellierung der Wasserflächen im Nationalpark Neusiedler See – Seewinkel (FEMOWINKEL). Endbericht von StartClim2021.G in StartClim2021: Handeln und Aktivieren, Auftraggeber: BMK, BMWFW, Klima- und Energiefonds, Land Oberösterreich.

Wien, im Dezember 2022

StartClim2021.G

Teilprojekt von StartClim2021

Projektleitung von StartClim:

Universität für Bodenkultur, Department für Wasser – Atmosphäre – Umwelt

Institut für Meteorologie und Klimatologie, Gregor-Mendel-Straße 33, 1190 Wien

www.startclim.at

StartClim2021 wurde aus Mitteln des BMK, BMWFW, Klima- und Energiefonds und dem Land Oberösterreich gefördert.

Inhaltsverzeichnis

Abbildungsverzeichnis

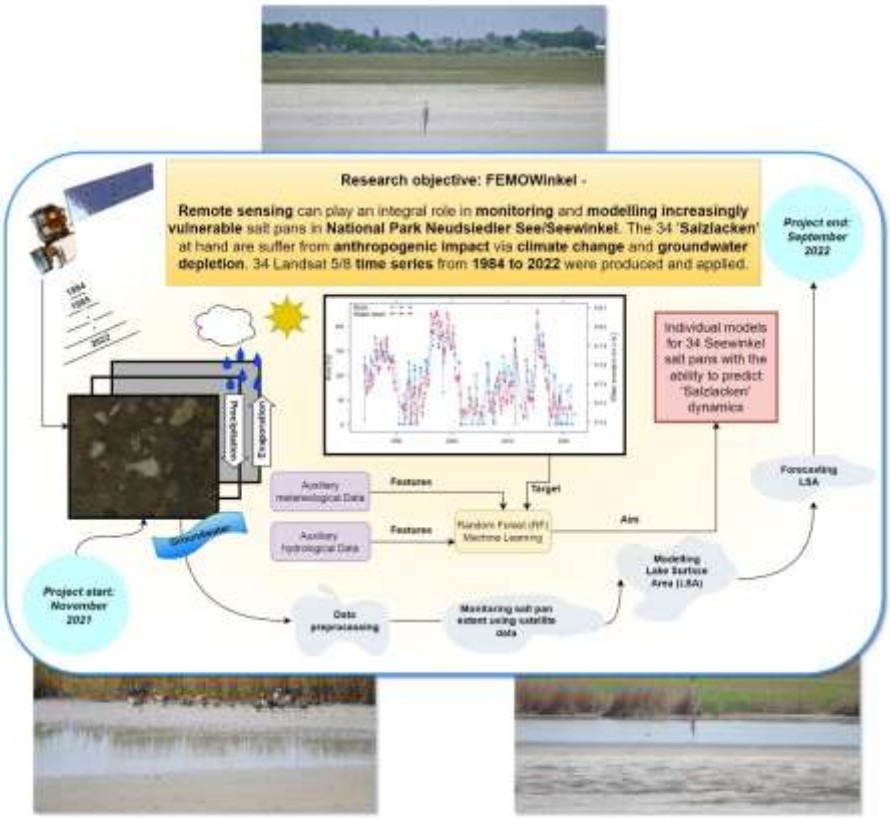
Tabellenverzeichnis

G-1 Kurzfassung

Feuchtgebiete sind von großer Relevanz für die Aufrechterhaltung der Biodiversität und erfüllen wichtige Funktionen in Bezug auf den Wasserhaushalt eines Gebietes. Die Salzlacken im Nationalpark Neusiedler See – Seewinkel stellen einzigartige Lebensräume für angepasste Tier- und Pflanzenarten dar. Ihr Erhalt beruht in großem Maße auf dem Wasserhaushalt des Gebiets. Die große Zahl von Lacken, die aufgrund ihrer geringen Tiefe oft nur sehr kurzfristig mit Wasser gefüllt sind und im Sommer teilweise austrocknen, erschwert das Monitoring mittels installierter Pegel. Satellitendaten stellen eine wichtige Informationsquelle dar, die flächenhaft konsistente Daten bereitstellen kann. Im Projekt FEMOWinkel wurden lange Zeitreihen auf der Basis von multispektralen Satellitendaten für das Monitoring der Lackenausdehnung verfügbar gemacht. In einem ersten Schritt wurde die Ausdehnung der Lacken basierend auf Zeitreihen der Landsat-Satelliten, die seit 1984 Daten liefern, abgeleitet.

Die abgeleiteten Zeitreihen wurden im zweiten Schritt mittels Methoden des maschinellen Lernens dazu verwendet, datengetriebene Modelle zu trainieren, die die Lackenausdehnung aufgrund von Klimadaten (Niederschlag, Temperatur, Verdunstung) und Grundwasserdaten modellieren können. Im Rahmen dieser datengetriebenen Modellierung ging es u.a. um die Fragestellung, ob es möglich ist, die sommerliche Austrocknung von Lacken bereits einige Monate im Voraus vorhersagen zu können, z.B. aufgrund zu geringer Niederschläge im vorhergehenden Winterhalbjahr. Die Ergebnisse zeigen, dass es für einen Großteil der Lacken möglich ist, unter Einbeziehung der Eingangsdaten für den Frühsommer kurzfristig vorhersagen zu können, ob eine Lacke im Sommer trockenfällt oder nicht. Bei längeren Vorhersagezeiträumen, z.B. wenn nur Daten bis zum Frühjahr vorhanden sind, sinkt die Wahrscheinlichkeit, eine korrekte Vorhersage zu treffen.

Die Projektergebnisse sind zum einen von Bedeutung für die wissenschaftliche Forschung, da die Vorhersage hydrologischer Zeitreihen mittels datengetriebener Methoden einen relativ jungen Forschungszweig darstellt. Zum anderen sind die Ergebnisse von Bedeutung für das Monitoring von Lacken ohne automatische Pegel, sowie für die Identifikation von besonders von Austrocknung gefährdeten Salzlacken.



G-2 Abstract

Wetlands are of great importance for biodiversity as a result of their habitat function. They also fulfil a vital role for the regional water cycle. The salt pans of the Austrian Neusiedler See - Seewinkel National Park - locally known as Salzlacken - are a unique habitat for specialised plant and animal species. Conservation of these vulnerable wetlands largely depends on their water balance and human water use. The large number of salt pans, which are also often inundated only for a short time, poses a challenge for operational monitoring. The monitoring is primarily carried out using in-situ gauges. Satellite remote sensing has the potential to bridge this gap by provisioning spatially consistent data. Within the FEMOWinkel project, long time series of water extent in the Seewinkel were retrieved from multi-spectral Landsat data reaching back in time as far as 1984.

In a second step, the retrieved time series of water surface area were used to train data-driven models to estimate water area and drying state of salt pans based on climate (e.g. precipitation, temperature, evaporation) and groundwater data. The data-driven models were used to address whether the drying state of a salt pan during summer can already be predicted some months earlier, e.g. due to missing recharge of salt pans and groundwater after low precipitation sums during the previous winter season. Results show that for a large number of wetlands a short-term forecast on the drying state is possible in early summer. With longer lead times the probability of making a correct forecast is decreasing.

The project results are of importance, on the one hand, for the scientific community as the data-driven modelling of hydrological processes is a relatively recent field. On the other hand, the results hold importance for the monitoring of salt pans with no in-situ gauges and the identification of wetlands which are particularly in danger of drying up.

G-3 Introduction

G-3.1 Background

Wetlands fulfil a multitude of ecosystem functions and provide services to the livelihoods of people and communities (Gardner et al., 2015). Regulatory services include water storage during dry periods and flood retention; provisioning services are mainly related to their habitat function for fish and other species used for food; cultural services are provided due to their important functions for tourism; supporting services include habitat provision for species, such as fish, amphibians and waterfowl, as well as important functions for biogeochemical cycles (e.g. Cheng & Basu, 2017). Supporting services are of special importance for the impacts of wetland loss on biodiversity. Despite these facts, wetlands are among the most endangered habitats. It is estimated that since the 1970s, 35% of wetlands have been lost (Convention on Wetlands, 2021), with losses reaching up to 90% in some regions, such as Europe and South Canada (Junk et al., 2012). Drainage and conversion to agricultural land, water abstraction for human use, pollution and alteration of environmental flows are among the chief causes for wetland degradation. Climate change is expected to exacerbate these impacts as extreme events, such as floods and droughts, are expected to become more frequent in many parts of the world (e.g. Mitchell, 2012).

G-3.2 Study area

The *Seewinkel* is located to the East of Lake Neusiedl near the Hungarian border in *Burgenland* (Austria) and is part of the Pannonian basin. Due to its complex geological and climatological structure the region is home to fens, lakes and, most important to this study, salt pans, locally known as *Salzlacken*. These shallow waterbodies, which rely on salt-rich groundwater, serve as a hub of biodiversity, being home to a large number of plant and animal species (Krachler et al., 2000), especially halophytes and migrant birds (Naturschutzbund Burgenland, 2012). The *Seewinkel* is affected by similar anthropogenic processes in respect to wetland degradation and ecological disconnectivity as seen in other regions of the world (Sharma et al., 2021). This is indicated by the declining number of waterbodies: in 1855, 139 *Salzlacken* were mapped, while in 2012, 59 exhibited an at least sufficient state (Naturschutzbund Burgenland, 2012). Decisions in favour of agricultural and town planning led to the construction of channels and wells in order to decrease the groundwater table, starting as early as 1828. Many others followed, most namely the *Einserkanal* in 1910, the *Hauptkanal*, *Zweierkanal* and the Xixsee-Kanal in the 1950s (Naturschutzbund Burgenland, 2012). The area around the salt pans, therefore, is heterogenous. Although some salt pans are surrounded by grassland, agricultural fields are normally in proximity (Dick et al., 1994). Forest are not abundant, only small tree-covered areas occur. A map of the currently existing *Salzlacken* is shown in Abb. G-1:. To understand and model surface water dynamics it is key to understand the complex hydrological regime, which in return is influenced by geological, meteorological, and anthropogenic factors. Annual precipitation in the region is ca. 590 mm, however, subject to high seasonality. Due to high summer temperatures, potential evaporation exceeds rainfall during summer (Soja et al., 2013). According to data from ERA5-Land (Muñoz-Sabater et al., 2021), the average 2 m air temperature is 10.8 °C with hot summers and moderately cold winters. The FEMOWINKEL project focuses on 34 *Salzlacken* that are filled by water at least periodically and can be observed from space using moderate-resolution (i.e. with a spatial resolution of several tens of metres) multi-spectral sensors. Before further explaining the key elements, some relevant literature is briefly summarised:

As satellite measurements and in-situ information complement each other, this study relies on, not only in-situ groundwater and precipitation measurements, but also on past research. The salt pans in the Seewinkel have been the object of several scientific and non-scientific publications: individually studying each salt pan, *Die Salzlacken des Seewinkels* (Naturschutzbund Burgenland, 2012) is a key resource. Hydrological, chemical, ecological & biological characteristics are described, finally conducting an evaluation for each wetland concerning its potential for restoration. From a hydrological point of view, only three salt pans are classified as grade 2(*Oberer Stinkersee, Obere Fürstenlacke, Mittlere Fürstenlacke*), seven as grade 3 (*Große Neubruchlacke, Westliche Fuchslochlacke, Lettengrube, Unterer Stinkersee, Südlicher Stinkersee, Kaschitzlacke*), while all others obtain a grade of 4 or worse on a five-grade scale ("1" signifying a very good condition, "5" a very bad condition). Degradation is, in all cases, linked to groundwater table decrease. As it was published in 2012 (and refers to the condition as it was in 2010), it cannot be considered an up-to-date publication but remains the most comprehensive source on the individual wetlands. Another, more detailed, study has been carried out by Krachler et al. (2000). Starting with the observation that the number of intact salt pans is declining, the authors focus on questions regarding the formation, the functioning and possible reasons of the phenomenon of 'dying-salt pans'. Despite this being a limno-

chemical analysis, they focus on finding mechanisms contributing to the decreasing ecological state of these small systems. Horváth et al. (2019) analysed airborne imagery acquired in 1959 and 2016 to map the abundance of salt pans in the project area. They found that the number of waterbodies in the region had declined from 110 to 30 between the two acquisitions. Along with the observed wetland loss, the number of invertebrate species in the remaining ponds declined.

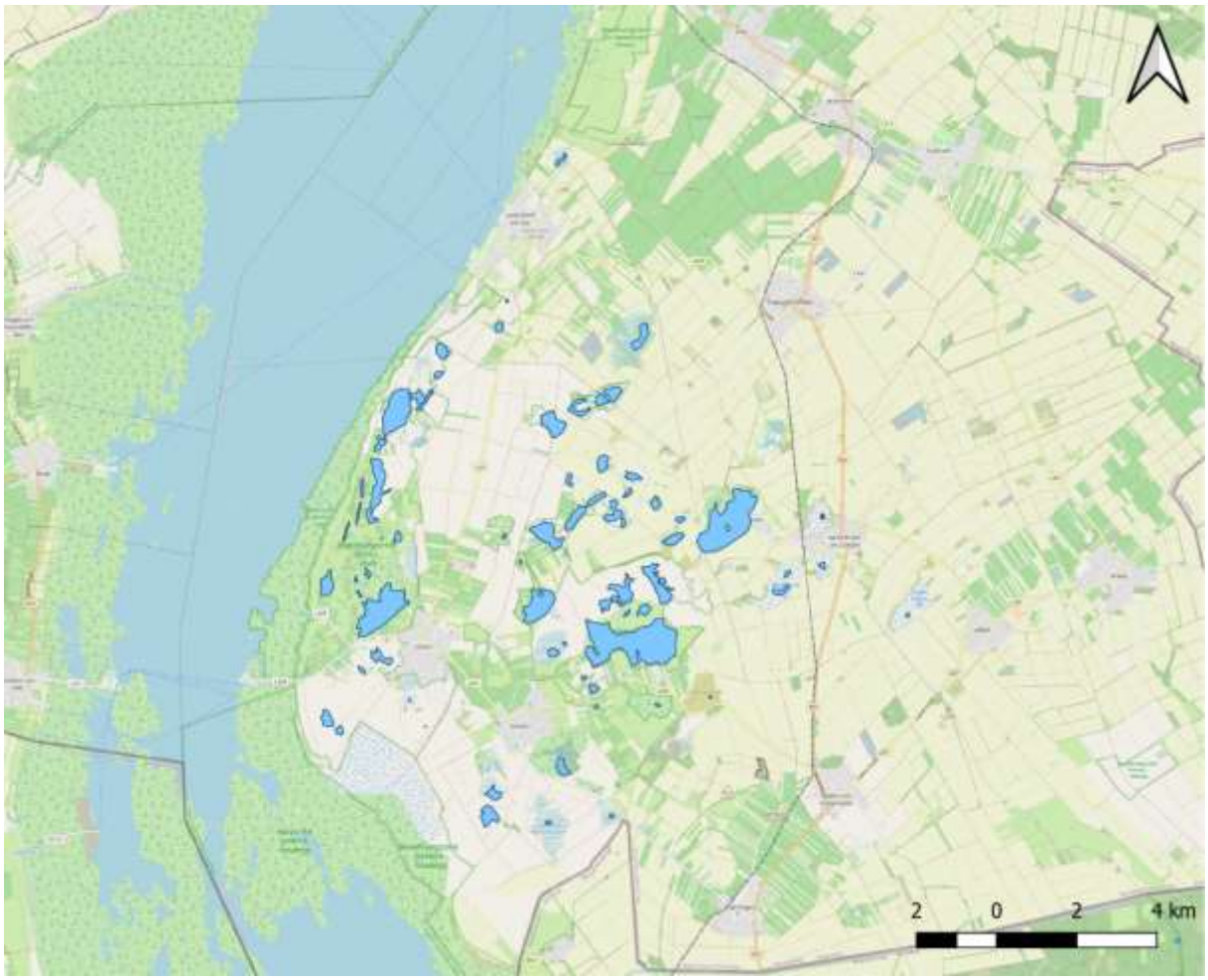


Abb. G-1: Salzlacken of the Seewinkel (source: M. Kuttner/Nationalpark Neusiedler See - Seewinkel). Base map and data from OpenStreetMap and OpenStreetMap Foundation (CC-BY-SA). © https://www.openstreetmap.org and contributors.

Distinguishing *Salzlacken* can be approached according to qualitative criteria, such as their chemical composition, geological/geomorphological formation or quantitative criteria, such as their dynamics in terms of water volume stored and the timing of inundation, the so-called wetland hydroperiod. The FEMOWINKEL project aims at understanding wetland dynamics in terms of water surface area. Therefore, the latter approach constitutes a sensible way of differentiation. Two general types (with two subtypes for the first category) of *Salzlacken* appear in the literature: perennial *Salzlacken* and summer-dry *Salzlacken* (Naturschutzbund Burgenland, 2012). Their main properties are summarised in Tab. G-1:.

Tab. G-1: Types of salt pans *(Naturschutzbund Burgenland, 2012)*.

Characteristic	Type 1A	Type 1B	Type 2
Water availability	Perennial	Perennial	Summer-dry

Supply	Precipitation, Groundwater	Precipitation, Groundwater, Externally fed	Precipitation, Groundwater
Water balance	Positive	Positive	Negative
Water level	Astatic	Astatic	Static
Occurence	Sparsely existent near lake-shore: 'Seerandlacke'	Sparsely existent	Common
Trend	Negative	Steady	Negative
Example case	*Untere Hölllacke, Unterer Stinkersee, Herrnsee*	*St. Martins Therme, Zicksee, Darscholacke*	*Lange Lacke, Große Neubruchlacke, Oberer Stinkersee*

Saline pan dynamics

The salt pan cycle has intensively been discussed in the literature (e.g. Chivas, 2007; Lowenstein & Hardie, 1985; Shaw & Bryant, 2011). A brief summary of the available literature is provided here, followed by a comparison of general natural salt pan dynamics to the specific dynamics of the *Seewinkel* salt pans. The first stage of the cycle is the flooding stage entailing a brackish lake. Inflow can be of twofold nature: due to local precipitation events or due to groundwater level being higher than the top of the impermeable layer. The second stage is the concentration stage leading to the development of a saline lake. The driving process in this stage is the evaporation of water, while dissolved substances remain in the shrinking waterbody. With further evaporation the salt concentration increases and salt deposition starts. When no water is present anymore, the desiccation stage is reached. This is the third stage of the salt pan cycle entailing a dry pan. This phase is the default stage of a salt pan. Periodic connection to the groundwater body is necessary for these ecosystems to persist (Lowenstein & Hardie, 1985). Bathymetry is of special importance in this respect. In contrast to deep lakes, where a small amount of rain might affect Lake Surface Area (LSA) only to a small extent, salt pans are shallow wetlands. Therefore, a single precipitation event could potentially result in considerable changes in LSA. Vice versa, a drought period can quickly lead to a shrinking of LSA, causing wetlands disconnection (Horváth et al., 2019) and fragmentation. This especially affects species reliant on local, immobile reproduction, such as arachnoids and amphibians (Naturschutzbund Burgenland, 2012). Furthermore, annual species, such as Große Salzmelde (*Suaeda pannonica*) and Kleine Salzmelde (*S. prostrata*) are more vulnerable to extreme wetness or dryness (Albert, 2013).

Dying salt pans

Salt pans are vulnerable to disturbances and longer-term trends in water supply and removal. Salt pans periodically fall dry when precipitation sums are low. During periods with higher precipitation these salt pans typically refill. The absence of rainfall or, more critically, the occurrence of low groundwater levels, however, cannot be compensated during wetter periods (Krachler et al., 2000). In the latter case, *drying from beneath* occurs, a process currently affecting many of the *Seewinkel* salt pans. It is caused by an increasing distance between the lake bottom and the groundwater table (threshold at 70 cm according to Naturschutzbund Burgenland (2007)), resulting in an interruption of capillary rise. As a result, no salt is transported to the wetland bottom weakening the impermeable layer. Additionally, precipitation is incrementally eluviating the impermeable layer from above, stopping lake formation while also transporting salts into deeper layers (Krachler et al., 2000).

G-3.3 State-of-the-art

G-3.3.1 *Monitoring of wetlands using earth observation data*

Wetland hydroperiod describes the duration and timing of water occurrence in a wetland and is of importance for characterising ecosystem functions, such as a wetland's habitat function (Foti et al., 2012). Its determination relies on precise measurements of water level and water extent. While *in-situ* water gauges provide accurate measurements at a high temporal resolution, installation and

maintenance of fixed water gauges are costly. Moreover, due to the point-based nature of water gauge measurements they only provide highly accurate information from their location, whereas heterogeneity in terms of bathymetry and vegetation can lead to considerable spatial variability within a wetland. Earth observation using satellite platforms can provide synoptic information about surface water extent over large areas (e.g. Hess et al., 2003; Pekel et al., 2016; Reschke et al., 2012). Satellite sensors can provide information about water extent and, to a limited extent, water surface elevation. The direct measurement of water surface elevation using radar altimeters, however, is confined to narrow tracks underneath the satellite overpasses and, hence, not available for most inland waterbodies and especially not for smaller ones due to the relatively large footprints of altimeters over land surfaces (Birkett, 2000; Schwatke et al., 2015). Hence, most studies have focused on the retrieval of surface water extent using optical (Klemas, 2013; Niculescu et al., 2020; Pekel et al., 2014; Vanderhoof & Lane, 2019) and microwave data (Hess et al., 2003; Prigent et al., 2007; Reschke et al., 2012). Especially data acquired using the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Landsat satellites have been used for the delineation of surface water bodies over large regions. Multispectral data from Landsat have been demonstrated to be of use for the distinction between water and land surfaces (Pekel et al., 2016) and have been used for the monitoring of inter-annual variations in the extent of small wetlands (Ogilvie et al., 2018; Vanderhoof et al., 2016). While vegetated areas are characterised by high reflectance in the near infrared (NIR) portion of the electromagnetic spectrum (ca. 700 to 800 nm), water surfaces typically absorb most of the incoming radiation at these wavelengths (Jensen, 2007). This makes NIR bands especially useful in the case of salt pans, which are often hard to distinguish from bare and sparsely vegetated areas (e.g., sand surfaces, lake sediments) in the visible wavelengths. Bare salt surfaces have been found to have much higher reflectivity in both visible and NIR bands than water surfaces (Safaee & Wang, 2020). Therefore, multispectral data are prime candidates for wetland monitoring in steppe regions, such as the Seewinkel. The data acquired by the Landsat 5 to 8 satellites furthermore covers a period of ca. 38 years, which makes them suitable for characterising inter-annual variations in wetland extent due to climate variability and land use change. Due to the global coverage of Landsat, large-scale operational surface water products have been generated, such as the Global Surface Water (GSW) product (Pekel et al., 2016) and the Dynamic Surface Water Extent (DSWE) product (Jones, 2019). The GSW product is globally available through Google Earth Engine as temporal composites and change maps as well as annual and monthly composites.

Cloud cover and cloud shadows are a limiting factor for the use of multispectral data. These effects are typically addressed during level-2 processing of the data. Cloud and cloud shadow masks are derived using a scene classification before atmospheric correction of the data. Commonly applied algorithms include the Land Surface Reflectance Code (Vermote et al., 2018) and Sen2Cor (Main-Knorn et al., 2017). More recent approaches have focused on the use of deep learning architectures for cloud and cloud shadow segmentation in multispectral images (e.g. Wieland et al., 2019).

G-3.3.2 *Wetlands modelling*

Wetlands and water bodies are typically modelled using physically based or conceptual models. Models, such as the Pothole Complex Hydrological Model (Liu & Schwartz, 2011), can model wetlands complexes and their size frequency distributions. However, they are often developed for certain types of wetlands, e.g., the North American Prairie Pothole Region in the case of Liu & Schwartz (2011), and specific amounts of data are necessary for their parameterization that are often not readily available. Data-driven modelling has emerged over the last decades as a flexible framework to bridge the gap between simple statistical models and physically based models. This progress has been facilitated by advances in the field of machine learning (ML). ML has been used to model ecosystem processes, such as evapotranspiration (Zhao et al., 2019), fire (Forkel et al., 2017) or discharge in karst environments (Xu et al., 2022).

In recent years, a number of studies have applied ML on predicting the water level of waterbodies (Wee et al., 2021), focusing on model interpretation, (e.g. Wu et al., 2022), model performance on a daily (e.g. Choi et al., 2020) or monthly (e.g. Hrnjica & Bonacci, 2019) basis. Water level is typically modelled based on *in-situ* measurements with a temporal resolution of 1 day or less, accounting for the short-term character of the hydrological cycle. Other studies, challenging alternative model constellations include (Chang et al., 2014; Ghorbani et al., 2018; Li et al., 2016; Nhu et al., 2020; Wang & Wang, 2020; Wen et al., 2019; Zhu et al., 2020). The feature spaces include hydrological and/or meteorological variables, while commonly struggling with proper estimation of peak values (e.g. Choi et al., 2020). Work on the modelling of LSA has been done on case studies in Australia (Soltani et al., 2021), Iran (Soltani

et al. 2021) and Central Minnesota (USA) (Delaney et al., 2022). For the former site, the authors applied Generalized Group Method of Data Handling (GGMDH) and stochastic modelling for a time series from between 2004 and 2019, while for the second one, SARIMA in combination with Deep Learning (DL) was used to model and predict remote-sensing derived LSA between 2001 and 2019. Modelling Lake Gregory (Australia) worked well with the GGMDH-model using either LSA, temperature and precipitation at time lag t-1 or a combination for short and long periods (2020-2060). Sumiya et al. (2020) and Albarqouni et al. (2022) employed correlation and regression analysis to relate LSA to hydro-meteorological variables. Another approach worth mentioning has been chosen by Daniel et al (2022). They introduce climate, LULCC (Land Use and Land Cover Change) and topography as features in a classification task aimed at estimating four different wetland permanence classes in the Prairie Pothole Region (PPR). Although struggling with relatively high error rates, permanence class membership was attributed to mainly climate or topography, depending on the study region.

G-3.4 Project goals

Based on the identified knowledge gaps on the Seewinkel Salzlacken ecosystems, the FEMOWINKEL project aims at addressing the following objectives:

1. Make use of multi-spectral Landsat imagery spanning the time period from 1984 to 2021 to propose a prototype for monitoring the extent of the salt pans in the region.

2. Apply ML methods to model water level and water extent in the salt pans based on climate and hydrological input data.

G-3.5 Project work breakdown structure

The work was organised in six work packages (WPs). The time planning for the WPs is shown in Abb. G-2:. The project was started in November 2021 and is scheduled to finalise in August 2022.

WP 1: Project management. This comprises all the activities necessary for a smooth implementation of the planned project activities as well as the participation in the two StartClim workshops.

WP 2: Data collection. Collection of all data necessary for the monitoring and modelling of the Salzlacken, such as earth observation data, climate data, groundwater levels and water levels. It also included interfacing with the Hydrographischer Dienst Burgenland.

WP 3: Derivation of wetland water extent. Analysis and classification of multi-spectral earth observation data into water and non-water areas for the time period 1984-2021.

WP 4: Typification of wetlands. Characterisation of salt pans in different categories according to their seasonal dynamics.

WP 5: Data-driven modelling of wetland water level and extent.

WP 6: Dissemination. Presentation of project results at StartClim workshops, scientific conferences and dissemination to stakeholders.

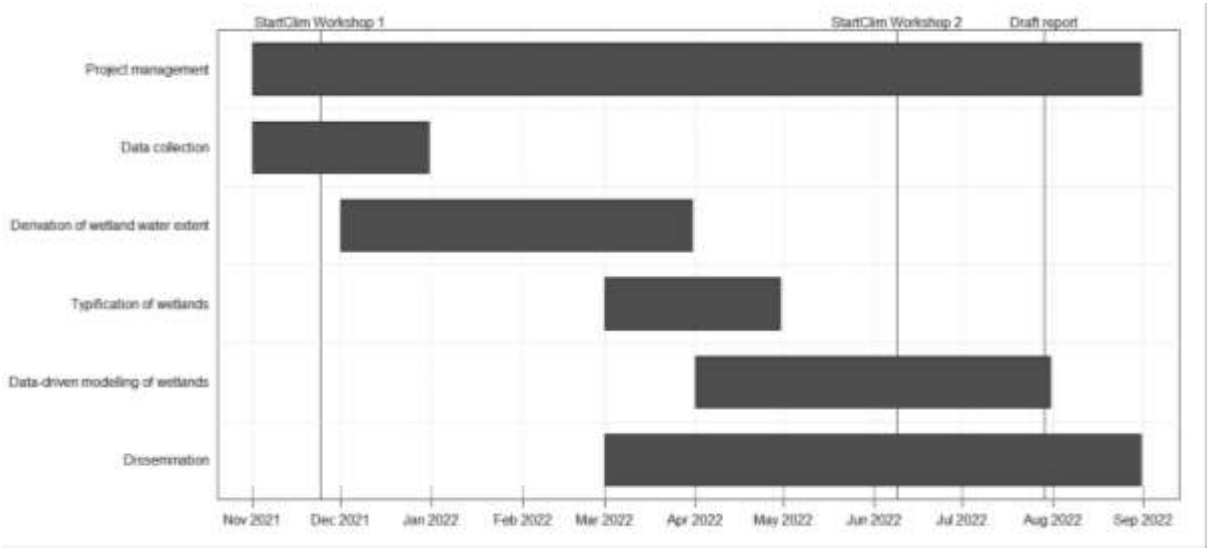


Abb. G-2: Time planning of the FEMOWINKEL work packages.

G-4 Methods

G-4.1 Mapping and monitoring of salt pans

G-4.1.1 *Datasets*

In the FEMOWINKEL project, we used multispectral data acquired by the Landsat 5 and Landsat 8 satellites. Landsat 5 was launched in March 1984 and carried the Thematic Mapper (TM) sensor which offered a spatial resolution of 30 m in six channels covering the visible, near and short-wave infrared spectra. The Landsat 8 mission was launched in February 2013 after the end of the Landsat 5 mission. Its main payload, the Operational Land Imager (OLI) has nine bands in the visible, near and shortwave infrared spectra with a spatial resolution of 30 m and a panchromatic channel with 15 m spatial resolution. The Thermal Infrared Sensor (TIRS) has a resolution of 100 m and two TIR channels. As we assigned a higher priority to a long time series than to a higher spatial resolution, we only used the 30 m OLI bands. Due to the end of the Landsat 5 mission in 2013 and the start of the Landsat 8 time series only after its commissioning phase was completed in March 2013 we have a gap of ca. one year from early 2012 to early 2013.

We refrained from using Landsat 7 data because of the scan-line corrector failure that occurred in May 2003 and affects the imagery acquired since that date. As a result, about 22% of data are missing during acquisition, however, this rate depends on the location of the target area within the footprint. Areas imaged at nadir are less affected than those at higher observation angles.

All data used in the project were available through the Google Earth Engine (GEE) (Gorelick et al., 2017). Landsat Collection 2 Level 2 data constitute a surface reflectance product that has been radiometrically corrected for atmospheric influences using the Land Surface Reflectance Code (LaSRC) algorithm (Vermote et al., 2018). An overview of the Landsat products used in the FEMOWINKEL project is shown in Tab. G-2:. We only used data acquired between April and October as we assumed that most dynamics would occur during the spring and summer months.

Tab. G-2: Landsat products used in the project.

Platform/ Sensor	Product	Time period	Bands used	Link
Landsat 5 TM	USGS Landsat 5 Level 2 Surface Reflectance, Collection 2, Tier 1	1984-2011	B1: 0.45-0.52 µm (blue) B2: 0.52-0.60 µm (green) B3: 0.63-0.69 µm (red) B4: 0.77-0.90 µm (NIR) B5: 1.55-1.75 µm (SWIR1) B6: 2.08-2.35 µm (SWIR2)	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LT05_C02_T1_L2
Landsat 8 OLI	USGS Level 2 Surface Reflectance, Collection 2, Tier 1	2013-2021	B2: 0.452-0.512 µm (blue) B3: 0.533-0.590 µm (green) B4: 0.636-0.673 µm (red) B5: 0.851-0.879 µm (NIR) B6: 1.566-1.651 µm (SWIR1) B7: 2.107-2.294 µm (SWIR2)	https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C02_T1_L2

A Landsat 8 image is shown in true colour as well as in false colour in Abb. G-3:. In the false-colour composite, water surfaces are visible as dark blue areas, whereas unvegetated agricultural areas are visible as red rectangles due to the higher shortwave infrared (SWIR) reflectance. Vegetated fields and natural vegetation appear in light green due to high NIR reflectance values.

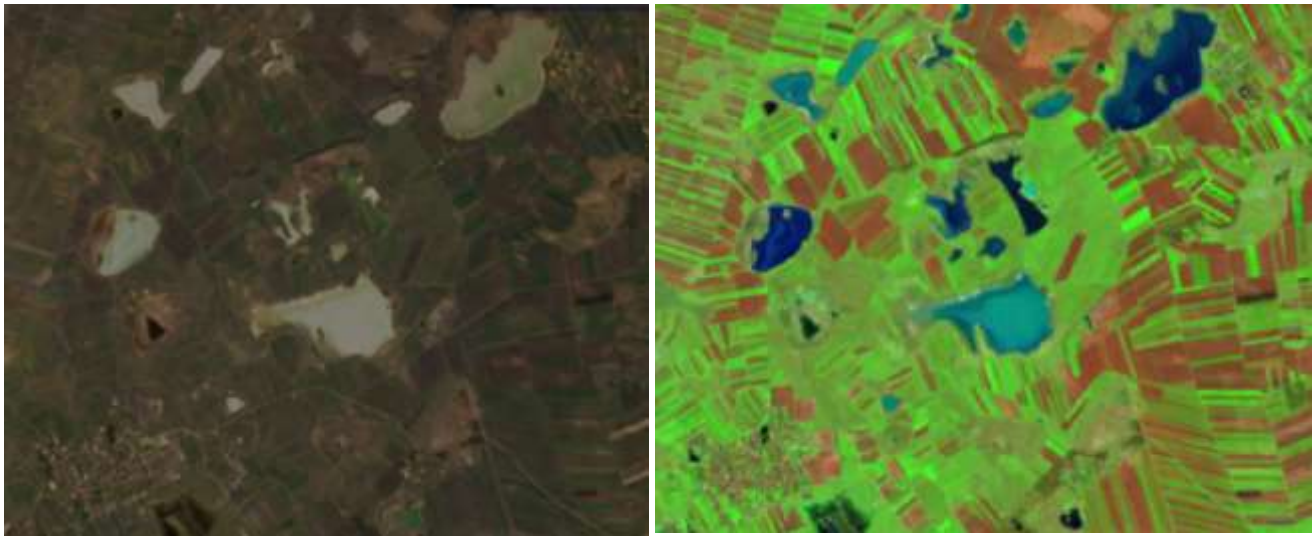


Abb. G-3: Landsat 8 OLI image as true-colour composite (left) and false-colour (Red: SWIR 1; Green: NIR; Blue: red channel) (right).

In addition to using the bands listed in Tab. G-2:, we derived the following spectral indices to be used as additional features in the classification:

- Normalised Difference Vegetation Index (Jensen, 2007) $NDVI = \frac{NIR-red}{NIR+red}$,
- Normalised Difference Water Index (McFeeters, 1996) $NDWI = \frac{green-NIR}{green+NIR}$,
- Modified Normalised Difference Water Index (H. Xu, 2006) $MNDWI = \frac{green-SWIR1}{green+SWIR1}$.

G-4.1.2 *Classification of salt pans*

For classification of the Landsat 5/8 surface reflectance products we selected the Random Forest (Breiman, 2001) classifier available in GEE[1]. We trained separate random forest (RF) models for each month between April and October in order to reflect the seasonality in reflectance. For this purpose monthly means were computed from the image time series for each spectral band and spectral index. As the spectral bands furthermore differ between the TM and OLI sensors we also trained separate models for data derived from Landsat 5 and Landsat 8. Furthermore, models were trained with and without the use of the indices NDVI, NDWI and MNDWI in order to evaluate the benefit of using these additional features.

Training polygons were selected over permanent water and land areas. 80% of the retrieved samples were used for training the models, the remaining 20% were retained for validating the classification results. The RF models were built with a number of 300 trees. The number of variables per split was set to the square root of the number of input features (i.e. 2 in the case of models without and 3 in the case of models with spectral indices).

We evaluated the models by computing overall accuracies from the test dataset. In order to obtain a more independent test dataset 10 Landsat 8 and 10 Landsat 5 scenes were randomly selected from the image stack and samples were drawn using the pixels retained for validation. This was done to make sure that no pixels were chosen for testing that were used for computing the monthly means that were used to construct the training features. We report the accuracy metrics for each combination of validation image, sensor and feature set (with and without normalised spectral indices).

[1] Available at https://developers.google.com/earth-engine/apidocs/ee-classifier-smilerandomforest

G-4.2 Data collection and exploratory data analysis

G-4.2.1 *Datasets*

Three types of datasets were used: in-situ data, reanalysis data & remote sensing-based data on wetland water area. Dataset selection was driven by the need to provide meteorological and hydrological data for the entire period covered by the remotely sensed dataset, i.e. from the beginning of 1984 until at least the end of 2021, rather than using datasets with a high spatial resolution but short temporal coverage. The bounding box coordinates used for data retrieval were, if needed, latitude 47.6° N to 47.9° N and 16.7° E to 17.0° E. If applied, anomalies were calculated by computing monthly average values over the entire period and then subtracting the respective monthly mean from the original value.

G-4.2.1.1 *In-situ datasets*

Information on groundwater level was accessed and downloaded from the Austrian eHyd portal[2] for measurement gauges located in *Seewinkel* (BUNDESMINISTERIUM FÜR NACHHALTIGKEIT UND TOURISMUS, 2018). Limitations regarding the covered time period narrowed down the number of available stations (from ca. 80 to six stations). Firstly, not all stations offered measurements back to 1984. Secondly, for up-to-date information these groundwater observation wells need to be equipped with automatic sensors, as otherwise the project would have needed to wait for manual evaluation of each sensor until the end of 2022. These data were provided by the *Hydrographischer Dienst Burgenland*. Following stations matched all criteria: S(station) 306043 St. Andrä am Zicksee, S 319418 Gols, S 316174 Neusiedl am See, S 305755 Frauenkirchen, S 305813 Illmitz and S 319426 Apetlon. The measurements are point-based and come in a monthly resolution. Additionally, all data are submitted to quality control (BUNDESMINISTERIUM FÜR NACHHALTIGKEIT UND TOURISMUS, 2018). For the yearly model, the monthly data were averaged in order to generate the yearly resolution.

Precipitation stations are scarcer in comparison to groundwater stations. Five stations were available within *Seewinkel*: S 122010 Illmitz, S 110585 Andau, S 110569 Apetlon, S 110551 Podersdorf am See and S 106724 Frauenkirchen. eHyd provided data until the end of 2018, while data from ZAMG extended the precipitation time series until 2022. Data are point measurements at a daily temporal resolution. The daily data were summed to monthly and annual scales.

Information on the recent water levels of some salt pans can be viewed and downloaded at Wasserportal Burgenland[3]. Longer records for 3 of the 34 salt pans, *Lange Lacke*, *Zicksee* and *Darscholacke*, were also provided by the *Hydrographischer Dienst Burgenland*. The measurements are point-based, although we were informed that not each station is located at the lowest point of the wetland (Abb. G-4:). This dataset still represents a reliable information-resource, again, being subject to quality control. Here, too, monthly data were averaged over a yearly long period for the classification task.

Abb. G-4: Water level gauges in *Lange Lacke* (left) and *Darscholacke* (right). Photos ©Schlaffer.

[2] Available at https://ehyd.gv.at/

[3] Available at https://wasser.bgld.gv.at/

G-4.2.1.2 *Reanalysis datasets*

ERA5-Land provided by Copernicus Climate Change Service offers a high-quality reanalysis product (Muñoz-Sabater et al., 2021). We used the monthly variables Total precipitation, 2m temperature and potential evaporation[4] on a 0.1° x 0.1° (ca, 9 x 9 km) grid from 1984 to 2021. Therefore, seven pixels were derived in total that were averaged in a subsequent step, resulting in a single time-series per variable. 2m temperature is given as monthly averages in °C and potential evaporation and total precipitation as monthly sums in mm.

G-4.2.1.3 *Calculation of Secondary Products*

The drought indices Standardized Precipitation Index (SPI) and Standardized Precipitation-Evapotranspiration Index (SPEI) were calculated in R (R Core Team, 2017) using the package SPEI (Santiago Beguería & Vicente-Serrano, 2017). The merit of these indices is that they provide an indication of short-term and long-term drought conditions, opening the opportunity for use as, e.g., groundwater drought indicator (Kumar et al., 2016). Both indices were based on ERA5-Land reanalysis data. The indices were calculated for integration periods of 3, 6, 12, 24 and 48 months. Hence, 10 more time series that represent potential predictors were produced. Addiationally, the number of days within a month with a maximum air temperature above 25 °C based was calculated based on ERA5-Land 2m air temperature.

G-4.2.1.4 *Remote-Sensing datasets*

For modelling, LSA as derived from Landsat using the approach described in section G-4.1 was used as target variable. The final temporal resolution of the dataset is depending on cloud cover. To homogenise the temporal resolution, the lake-wise time series were averaged to a monthly interval. As some months turned out to have no information assigned at all, the time series were linearly interpolated by deploying the package scipy-interpolate[5] to produce additional and consistent data. Anomalies were calculated following the procedure described above.

G-4.2.2 Exploratory analysis

G-4.2.2.1 *Time series analysis*

Visual interpretation of the time series along with analysing trends and seasonal dynamics formed the basis for further analysis. Trend detection was carried out by applying the non-parametric Mann-Kendall test (Kamal, Neel; Pachauri, 2018) for each time series.

G-4.2.2.2 *Correlation analysis*

Three approaches are most commonly used in correlation analysis: *Pearson correlation coefficient* (R), *Spearman's rank correlation* (ρ) and *Mutual Information* (MI). For this study, the former was chosen as it fulfils the needs of a preliminary understanding between the variables. The calculations were performed using the Python package *pandas*[6]. In some cases, and for reasons of completeness (in order to capture non-linearity), Spearman's ρ was calculated as an additional metric.

G-4.3 Data-driven modelling of wetlands water level and extent

Monthly WL/LSA time series enable different types of model setups that in return open the possibility of exploring different variable links and performance optimization. But which model fits best to the case at hand? As WL/LSA was not introduced to the feature space, different hydrological and meteorological features were related to WL and LSA data. Each serving its own purpose, three distinct experiments were set up based on the RF approach (Breiman, 2001):

---

[4] The description of the variables can be found here:

https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=overview

[5] Available at https://docs.scipy.org/doc/scipy/tutorial/interpolate.html

[6] Available at https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.corr.html

1. WL was modelled using a RF regression approach. WL was used in this context as the in-situ based measurements available with a high temporal resolution for a limited number of wetlands were assumed to have a higher accuracy than the estimates of LSA based on remote sensing. It was assumed that the higher accuracy data for training and testing would increase the robustness of fitted models and estimates of feature importance.
2. LSA at a monthly scale was modelled using a RF regression approach with the goal of capturing the full dynamic range of each salt pan individually. The feature selection is the same as in the water level case.
3. In order to test the capability of ML methods for forecasting whether a wetland will fall dry during a given year, a RF classification approach was selected. As a target variable, for each year it was determined if LSA in a wetland reached zero. At this point it was assumed that the respective wetland dried out completely.

For modelling, the *scikit-learn* package, a free python machine-learning library, was selected[7]. It provides not only the RF-algorithm itself, but also various functionalities enabling explainability. Anomalies computed as described in G-4.2.1.1 were used for the modelling in order to remove the seasonality in the input time series.

The input data are time series exhibiting autocorrelation. Thus an iid-(independent identically distributed) assumption is not met. Possible strategies have been proposed to evaluate model performance of algorithms applied to time series (Cerqueira et al., 2020). Two cross-validation (CV) schemes would fit to this model: 1) Blocked Leave-One-Out CV (LOO-CV), and 2) prequential CV (NTS-CV). LOO-CV uses all parts of the data for training and testing in each cycle (fold). In the classification model, e.g., yearly data is being investigated, therefore, the test set includes one year, and the training set includes the other 36 years in respect to the whole timespan. Meanwhile, the NTS-CV uses a growing window of blocks for training and always test on a block later in time. We decided to use LOO-CV for the purposes of this study. The data splits for modelling both classification and regression are based on a 75% - 25% allocation. Finally, the processing chain shown in Abb. G-5: was used for the modelling, both for the regression and the classification task.

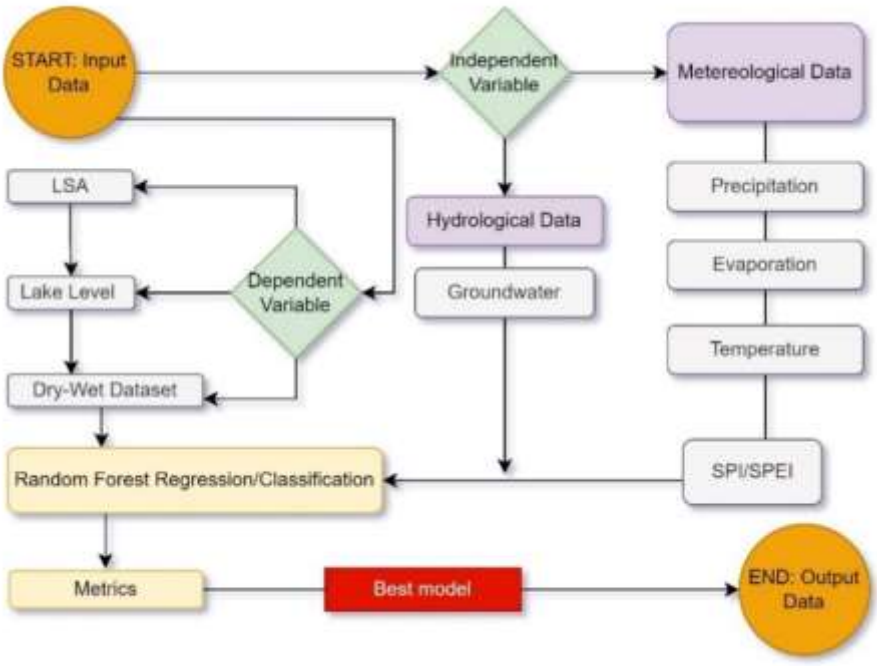


Abb. G-5: Flowchart of RF modelling.

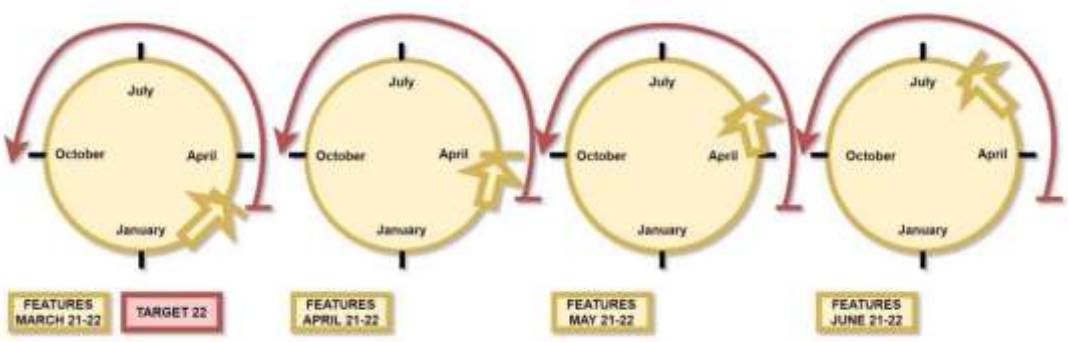
[7] Available at https://scikit-learn.org/stable/

Each model's hyperparameters were tuned using the scikit-learn package RandomizedSearchCV[8] function that is building on the concepts of Bergstra & Bengio (2012). Its main advantage is focus on parameter budgeting and adjustability while aiming at finding optimal parameters. Additionally, it poses an option for preventing overfitting that turned out to be an issue in this study. Models dealing with small amounts of data tend to overfit the data. This tendency was addressed by trimming the decision trees (key parameters: n_estimators, max_features, max_depth and min_samples_leaf) in scope of execution the aforementioned method. Apart from filtering, feature importance was applied by employing the scikit-learn package feature_importances[9] , RFE[10] (Recursive Feature Elimination) and RFECV[11]. Feature importance as a concept is based on the mean and standard deviation of accumulation of the impurity decrease within each tree and has proved to be a useful concept for RF classification (Saarela & Jauhiainen, 2021) and RF regression tasks (Grömping, 2015). As the classification task uses fewer and more basic predictors compared to both regression tasks and has a simpler setup in general, it will described first.

G-4.3.1 *Classification*

G-4.3.1.1 Model Setup

As seen in section G-3.2, salt pans are an extreme environment in terms of variability. This insight can be used for modelling, as the remotely sensed data explores LSA, and therefore inherits the ability to separate between these two opposed salt pan states . The null hypothesis was picked as: 'the salt pan will not dry out completely during summer'. These conditions were translated into the binary classes "the wetland falls dry in year X" (coded as 0) and "the wetland does not dry out in year X" (coded as 1). Besides, it was assumed that forecasting of wetland drying during a year gets more challenging with increasing lead time, i.e., the model is designed to lack information acquired during the summer months whose conditions heavily influence wetland water balance. Therefore, we decided to sequentially add information on the forthcoming month, resulting in four models: one each for March, April, May and June (Abb. G-6:). Each model includes information from the preceding 12 months. For example, the model for the month March 2022 uses information dating back to April 2021, the model for month April 2022 consists of information dating back to May 2021 and so on. For the testing period from 2011-2021, training data from the remaining period from 1984-2010 was chosen, resulting in 39 points of data. Starting the training data-window in autumn when evaporation declines and precipitation starts to gain significance or adding one without removing the last one pose other possible setups. The final setup is illustrated in Abb. G-6:.



8Available at

https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html

[9] Available at https://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html

[10] Available at https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html

[11] Available at

https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFECV.html#sklearn.feature_selection.RFECV

Abb. G-6: Classification experiment design.

For modelling the binary classification RF classifier was used. The estimator is included inside the scikit-learn Python package[12].

G-4.3.1.2 *Predictors*

Drawing from the considerations above with the additional restraining factor of data availability in mind (anthropogenic information), the predictors in Tab. G-3: constitute the model.

Tab. G-3: Predictors used for classification tasks.

	Temporal resolution	Spatial resolution	Source
Groundwater level anomalies	Monthly	Point-based	eHyd, Wasserportal Burgenland
Precipitation anomalies	"	Point-based	ehyd, ZAMG
SPI 6/24	"	0.1°x0.1°	ERA-5 Land
Evaporation anomalies	"	"	"
SPEI (not used for modelling)	"	"	"
Temperature anomalies	"	"	"
Number of days above 25 °C	"	"	"

G-4.3.1.3 *Evaluation*

As a performance metric for the binary classification we chose the F1-Score (Goutte & Gaussier, 2005), as it is commonly used for such problems. It is calculated as the harmonic mean between precision and recall. Furthermore, confusion matrices and other visualizations of the classification results were analysed.

G-4.3.2 *Regression Water Level (WL)/ Lake Surface Area (LSA)*

G-4.3.2.1 Model Setup

The regression model is aimed at predicting the variables WL and LSA within the same setup, although following a certain order. In a first step, the model with the target WL was tested and interpreted. In the second step, models were trained on LSA, which was assumed to be less precise, however, covering a larger number of wetlands. As the LSA was aggregated to a monthly interval modelling WL used the same monthly resolution to ensure comparability. The monthly WL/LSA anomalies were used as target variables. The models were trained using data from May 1984 to Dec. 2012 while the remaining data were used for model testing.

G-4.3.2.2 *Predictors*

The features were based on the understanding that meteorological and hydrological conditions not only from the currently tested month but also from the previous months are of importance. As the interpretability increases when adding the currently tested month, it was decided include such

[12] Availabe at:

https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

information despite decreasing the predictive utility. As the SPI/SPEI 6 and SPI/SPEI 24 (SPEI was not used for modelling, as evaporation data was picked as a separate predictor and the SPI is simpler, while the graph is showing a very similar behaviour (see G-5.2.1.2)) are covering information on drought conditions from 6 months to 2 years into the past, the other variables were expected to cover the remaining impact on the salt pans. As a conservative estimate, antecedent information of up to four months was included (Tab. G-4:), although 2-3 months were expected to be sufficient already.

Tab. G-4: Features used for the regression task in addition to variables listed in Tab. G-3:.

	Temporal resolution	Spatial resolution	Source
Groundwater anomalies averaged over 2,3,4-month	Monthly	Point-based	eHyd, Wasserportal Burgenland
Precipitation anomalies averaged 2,3,4-month	Monthly	Point-based	eHyd, ZAMG
Evaporation anomalies averaged 2,3,4-month	Monthly	0.1°x0.1°	ERA-5 Land

G-4.3.2.3*Evaluation*

As a performance metric for the regression model, we chose the Root Mean Square Error (RMSE). As RMSE is given in the same units as the variable of interest, the obtained values were normalised using the average area of each wetland to ensure comparability between the salt pans.

G-5 Results and Discussion

G-5.1 Mapping and monitoring of salt pans

Using the RF classifier, high to very high classification accuracies could be obtained for the retrieved water and non-water pixels for randomly selected Landsat 8 (Tab. G-5:) and Landsat 5 (Tab. G-6:) scenes. It is noteworthy that the models built using the indices NDVI, NDWI and MNDWI performed consistently better than the models that only use the information contained in the spectral bands. This is likely owing due to the clear separation between water and non-water pixels in the area of interest and its comparatively small size. A further aspect is the relative homogeneity of the area in terms of land-cover classes. Non-water pixels mainly belong to agricultural areas, grassland, and built-up areas, whereas forests or topographically challenging terrain are virtually absent. While these high accuracies indicate a good performance for the Seewinkel it is likely that the models will perform significantly worse when being transferred to other study areas due to the low diversity of the training dataset.

Tab. G-5: Overall accuracies for randomly selected Landsat 8 scenes.

Image date	With spectral indices	Without indices
2021-06-24	1.0000	1.0000
2019-10-25	1.0000	0.9996
2013-10-01	1.0000	0.9996
2014-05-20	1.0000	1.0000
2016-03-31	1.0000	0.9996
2018-04-29	1.0000	1.0000
2015-09-21	0.9975	0.9971
2018-07-27	1.0000	1.0000
2015-07-26	0.9879	0.9833
2013-07-29	1.0000	1.0000

Tab. G-6: Overall accuracies for randomly selected Landsat 5 scenes.

Image date	With spectral indices	Without indices
1995-10-23	1.0000	1.0000
1992-07-19	0.9747	0.9319
1993-09-24	0.9924	1.0000
1992-05-16	1.0000	1.0000
2006-06-15	1.0000	1.0000
1998-08-12	1.0000	1.0000
1985-07-16	1.0000	0.9933
1994-04-04	1.0000	1.0000
1990-05-11	1.0000	1.0000

1994-07-25	1.0000	1.0000

The comparison of the retrieved salt pan water extents with water levels measured in situ showed some limitations of the obtained results for the three Salzlacken where these data were available. In the case of the *Lange Lacke* (Abb. G-7:), water area shows high sensitivity to changes in water level as indicated by the high correlation between the two variables (Pearson r = 0.85, Spearman's ρ = 0.92). Both time series display distinct time periods with high water level/large water area, e.g., between 1996 and 2000 and around 2015, as well as periods with low water level/decreased water area, e.g. during the early 1990s, between ca. 2002 and 2010 and from 2017 until 2021.

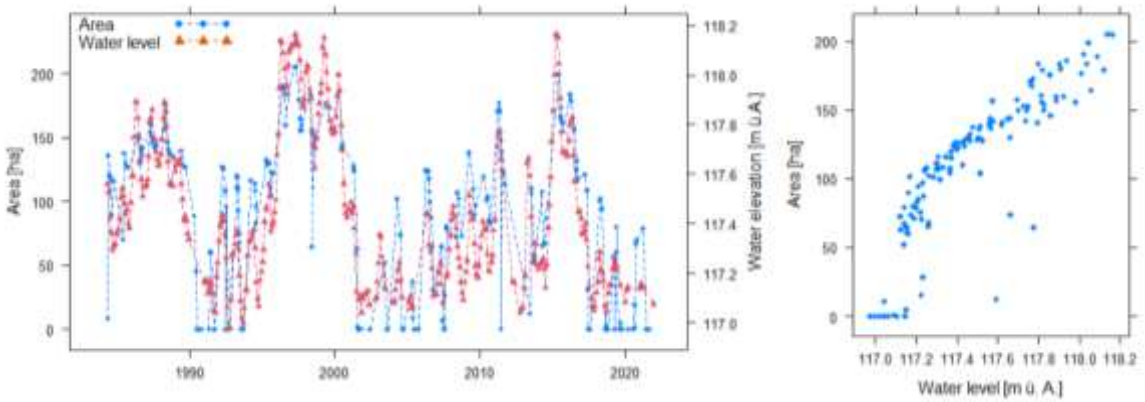


Abb. G-7: Time series of water area and water level (left) and scatterplot of the two variables for *Lange Lacke*.

In the case of the *St. Andräer Zicksee* (Abb. G-8:), the relationship between the two variables is entirely different. While in-situ water level measurements show dynamics similar to *Lange Lacke*, water surface area is more stable with the exception of single isolated points where area almost is 0. These are likely to be artefacts caused by incomplete cloud masking leading to an underestimation of the true water extent. However, during some of the drier periods a slightly lower water surface area can be observed, e.g. from 2018 onwards. Correlation coefficients are r = 0.42 and ρ = 0.56.

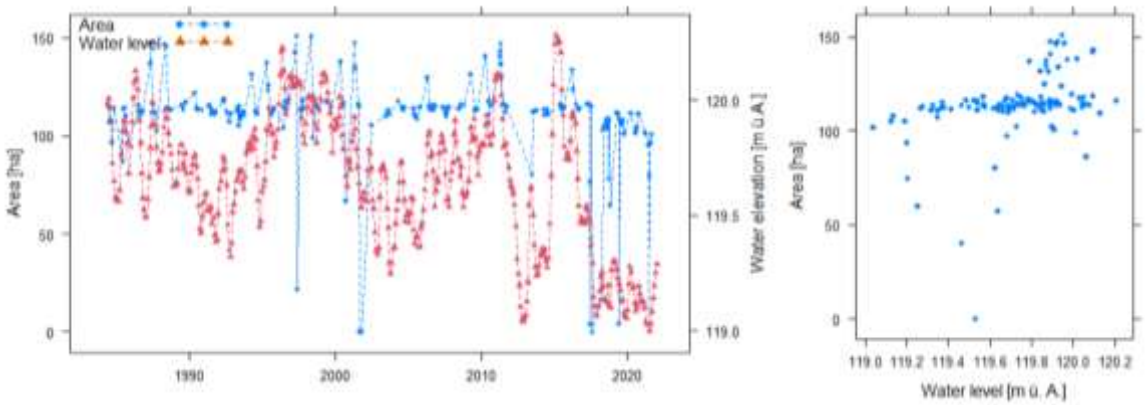


Abb. G-8: Time series of water area and water level (left) and scatterplot of the two variables for *St. Andräer Zicksee*.

At the *Darscholacke* (Abb. G-9:), both of the aforementioned aspects can be observed. On the one hand, water area follows the general dynamics of drier and wetter periods in the water level time series. However, the amplitude of water area is considerably lower. Nevertheless, the obtained correlation coefficients are moderately high (Pearson r = 0.52, Spearman's ρ = 0.73) indicating a fairly strong sensitivity of water area to changes in water level. On the other hand, outliers occur throughout the time series, which are likely again the result of incomplete cloud masking.

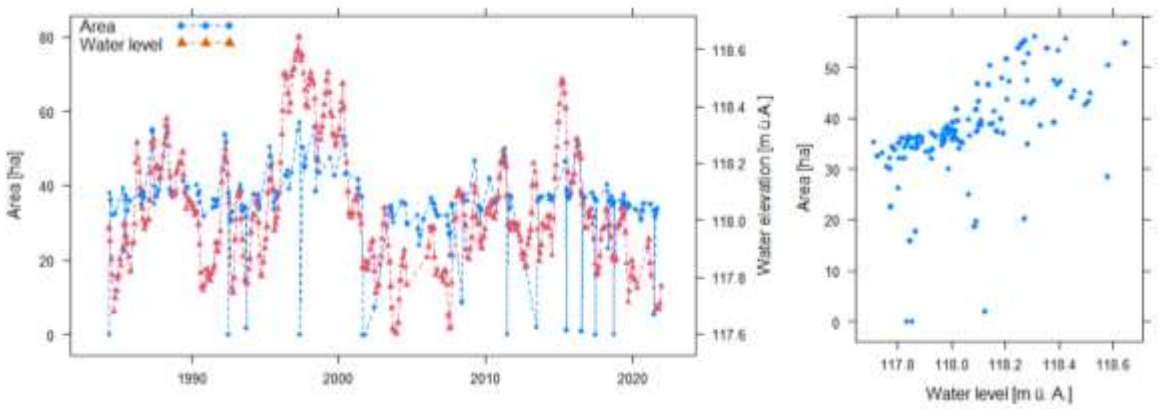


Abb. G-9: Time series of water area and water level (left) and scatterplot of the two variables for Darscholacke.

G-5.2 Exploratory Data Analysis

G-5.2.1 *Intravariable Analysis*

G-5.2.1.1 *Groundwater*

The six groundwater time series display a seasonal pattern with an increase from late-winter to early-summer and decrease during mid-late summer and autumn (Abb. G-10:). The station Neusiedl am See forms an exception, likely due to being influenced by the lake's water level. The *Mann-Kendall test* reveals significant decreasing trends (fitted lines shown in Abb. G-10:) for all stations except station 316174 for which the null hypothesis ('there is no significant trend') could not be rejected. All plots are showing high periods around 1970 (where data are available), 1995 and 2015, while around 1990 and 2004 low levels were measured. Depending on the point of measurement, different overall levels can be observed (e.g., the station 319426 is exhibiting lower mean values than station 319418, partly disclosing the structuring of the groundwater catchment).

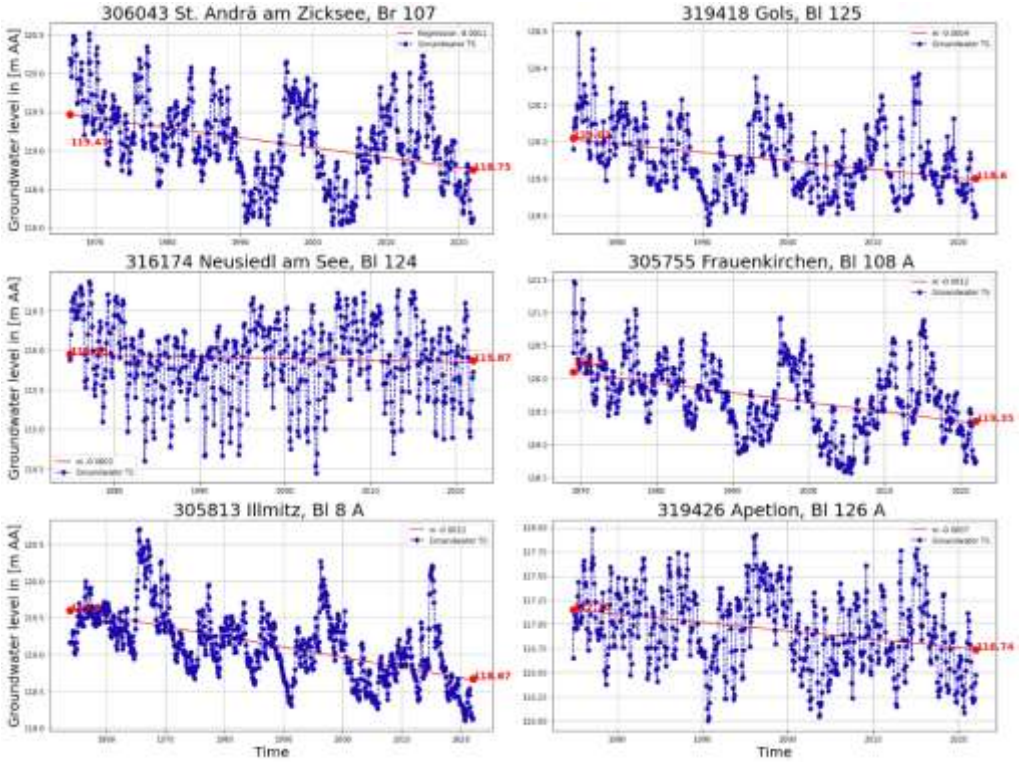


Abb. G-10: Time series of different groundwater stations (blue) with linear regression (red) for entire available periods.

G-5.2.1.2 *Drought indices*

The calculated drought indices SPI 6/24 and SPEI 6/24 are shown in Abb. G-11:. Depending on these integration times, the signal inherits a higher or lower sensitivity to periods of increased or decreased dryness/wetness. Drought episodes can be spotted in the beginning 1980s, the 1990s and around 2004. The SPEI 24 is additionally showing drought conditions starting in 2018. Wet episodes occurred during the late 1990s, around 2010 and 2015.

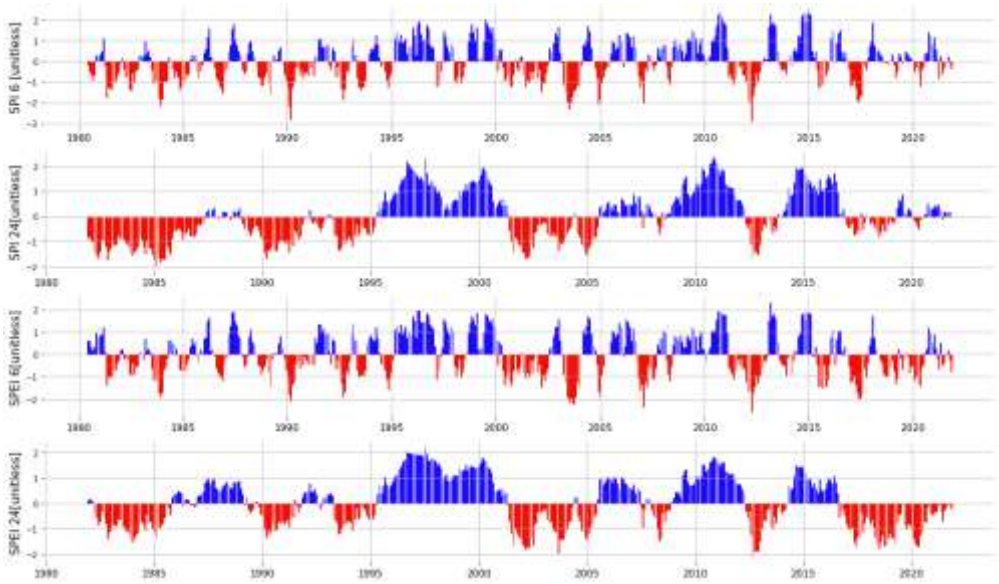


Abb. G-11: SPI and SPEI for integration periods of 6 and 24 months based on precipitation and evaporation data from ERA5-Land.

G-5.2.2 *Intervariable Analysis*

The matrix of *Pearson correlation coefficients* (Abb. G-12:) between the water level measured at *Lange Lacke (319608), Zicksee (319491)* and *Darscholacke (319525)* with all available variables described in G-4.2.2 reveals a moderate to strong correlation between water level and groundwater measurements (0.55, 0.43, 0.50, resp.). The respective coefficients between the three salt pans are very similar, yet the artificial salt pans *Darscholacke* and *Zicksee* show a weaker relationship in respect to groundwater. Contrary to the expectation, no relationship between precipitation (monthly), temperature (monthly), potential evaporation (monthly) and WL (monthly) can be revealed through straightforward correlation. Performing correlation on a daily basis and for various time lags has been carried out, while not revealing further insight. Yet, the drought indices exhibit a moderate correlation with wetland water level, anticipating the relationships that are further explored in section G-5.3.1.

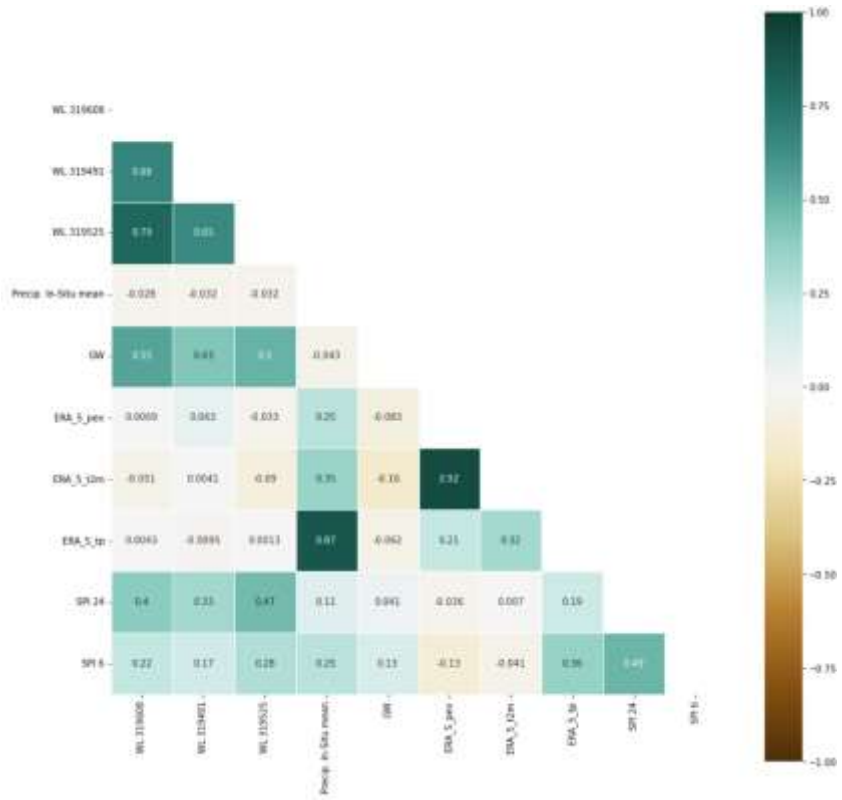


Abb. G-12: Correlation matrix for water level (*Lange Lacke (319608), Zicksee (319491)* and *Darscholacke (319525)*) and input variables. These include: Precipi. In-Situ mean (averaged precipitation over all in-situ stations), GW (groundwater), ERA_5_pev (potential evaporation), ERA_5_t2m (2m temperature), ERA_5_tp (total precipitation), SPI 24, SPI 6.

The correlation coefficients between Lake Surface Area (LSA) of all 34 water bodies with the selected hydro-meteorological variables are displayed in Abb. G-13: (again, similar results for Spearman's ρ were found). At first glance it shows that *Kiesgrube, St. Martins Therme 1 and St. Martins Therme 2* stand out as they show no or negative correlation regarding the other water bodies and are strongly correlated with each other. The relationship between the LSA of *Lange Lacke, Zicksee* and *Darscholacke*, on the one hand, and groundwater, on the other hand, is in accordance to results of the correlation analysis for WL and groundwater. Moreover, the correlation between precipitation and the salt pans is, again, weak. Commonly, evaporation and temperature are exhibiting a weak to medium negative correlation. The SPI 24 is showing a stronger correlation to the salt pans compared to the SPI 6.

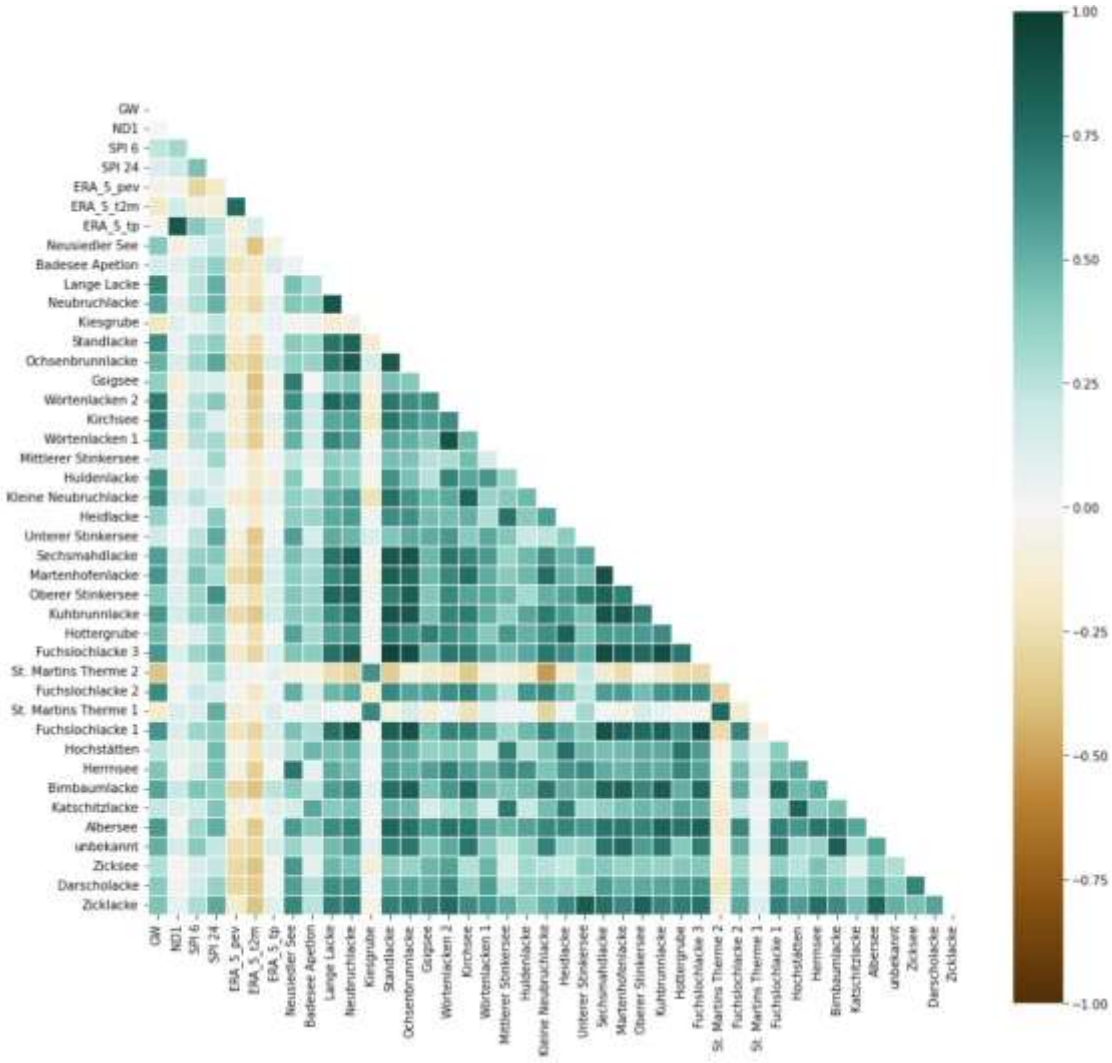


Abb. G-13: Correlation matrix for lake surface area (LSA) for all 34 salt pans and various other variables. These include: GW (groundwater), ND1(averaged precipitation over all in-situ stations), SPI 6, SPI 24, ERA_5_pev (potential evaporation), ERA_5_t2m (2m temperature), ERA_5_tp (total precipitation),

G-5.3 Data-driven modelling of wetland water level and extent

The classification targeting LSA scored a maximum (mean) F1-Score of 0.81 (model: June) with heterogeneous performance regarding the different salt pans. This statement holds true for both the regression models targeting WL and LSA (concerning the test set). Although the models generalize the ground-conditions by spatial averaging (which, as it turns out, does not pose a problem for model performance), the main factor attributing uncertainty is thought of being the varying level of anthropogenic forcing on the micro-ecosystems, as has been argued in section G-3.2. Another reason might be temporal asynchronisms and temporal averaging preventing adequate model timing (although not suggested by the correlation analysis). On another note, and as discussed in G-4.3, overfitting needed to be handled by trimming the decision trees for both regression and classification. Hyperparameters were tuned using cross-validation as implemented in *RandomizedGridCV*.

G-5.3.1 *Modelling Water Level (WL) via a Regression Approach*

The models including groundwater performed better at predicting WL than the models without this information (Tab. G-7:). This holds true for all the three salt pans where WL data were available: *Zicksee*, *Lange Lacke* and *Darscholacke*. The testing error for *Zicksee* is higher (0.39, 0.42 resp.) compared to *Lange Lacke (0.15, 0.24 resp.)* and *Darscholacke* (0.16, 0.23 rep.). Although unexpected in this extent, the large increase of test error compared to the training error for *Zicksee* likely hints at overfitting.

Tab. G-7: RMSE for regression on water levels of *Zicksee*, *Lange Lacke* and *Darscholacke* for the test-period. Maximum water level amplitude of all three salt pans is 1.25m.

Model	RMSE with groundwater	RMSE w/o groundwater
RMSE Zicksee [m] - Test	0.39	0.42
RMSE Lange Lacke [m] - Test	0.15	0.24
RMSE Darscholacke [m] -Test	0.16	0.23
RMSE Zicksee [m] - Train	0.06	0.09
RMSE Lange Lacke [m] - Train	0.11	0.16
RMSE Darscholacke [m] - Train	0.09	0.13

With a maximal amplitude of about 1.25 m of WL change for all lakes, the findings above indicate a model that, relative to that amplitude, performs adequately for *Zicksee* and rather well for *Lange Lacke* and *Darscholacke*. Zicksee and Darscholacke are both artificially fed, resulting in a larger susceptibility to error, as can be exemplified by looking at the large bias in the periods beginning in 2017 and 2020, respectively. Still, the correlation between the estimated and measured values indicates that the model results are closely related to the ground truth, as presented in Tab. G-8:. On another note, variance seems to be connected to short-term changes and bias to long-term dynamics while, in this model, both seem to be driven by changes in the groundwater body.

Tab. G-8: Pearson's r correlation coefficient between measured water level and estimated water level for all six model setups.

Pearson's r	Zicksee GW	Zicksee	Lange Lacke GW	Lange Lacke	Darscholacke GW	Darscholacke
WL - Overall	0.77	0.68	0.93	0.84	0.9	0.81
WL - Train	0.96	0.89	0.94	0.89	0.93	0.89
WL - Test	0.84	0.66	0.89	0.7	0.77	0.49

Lange Lacke displays a pronounced increase in WL from 1990 to 1997 (Abb. G-14:). This increase and the subsequent decrease are captured well by the model, although peak values are underestimated. A similar pattern can be spotted in the test period. With respect to the most recent years model estimates slightly overestimate the measured WL, especially when GW level is not used. The peak around mid-2015 is not portrayed in its full extend. Overall, as the scatterplot indicates, the training period is captured more nicely (Pearson's r: 0.94, 0.89 resp.), while the dynamic in the test period is captured but displaying a larger spread (0.89, 0.7 resp.).

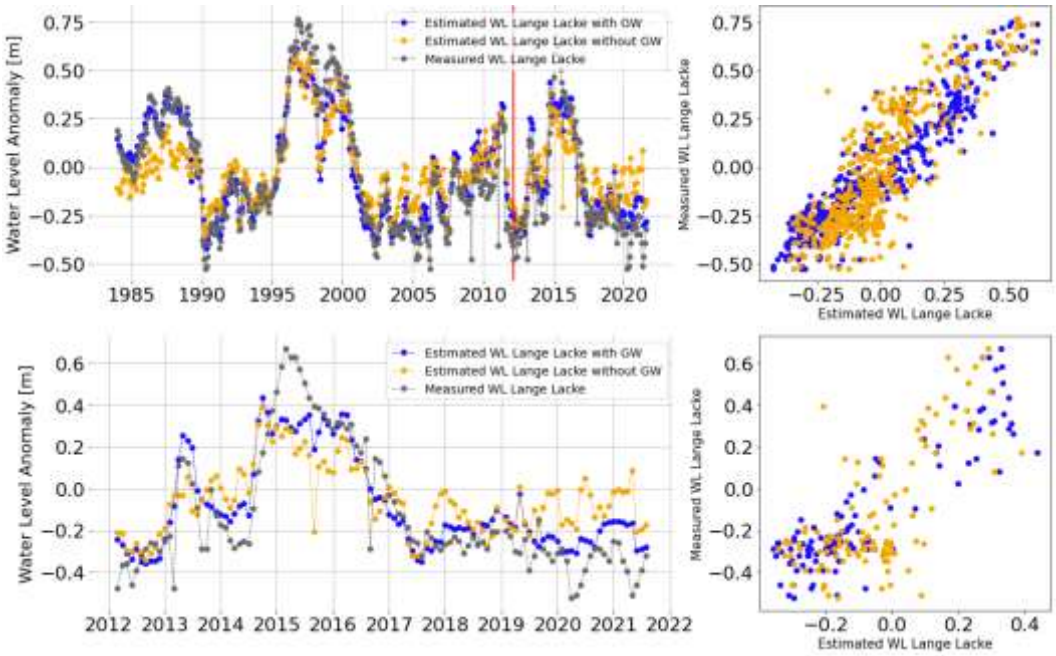


Abb. G-14: Regression result for *Lange Lacke*. The red line shows the separation between training (before) and test period (after).

The dynamic range of *Zicksee* (Abb. G-15:) is less pronounced for the training period compared to the testing period hindering an estimation of the systematically low values from 2012 onwards. Here, certain patterns relating to lower overall water levels are failed to be recognized, introducing a visible bias. Moreover, the model struggles to replicate variability, resulting in a lower overall correlation of 0.77 and 0.68 in respect to the model without groundwater contribution. The testing period correlation is much better when taking groundwater into account (0.84). This turns out to be a valid observation for all three salt pans.

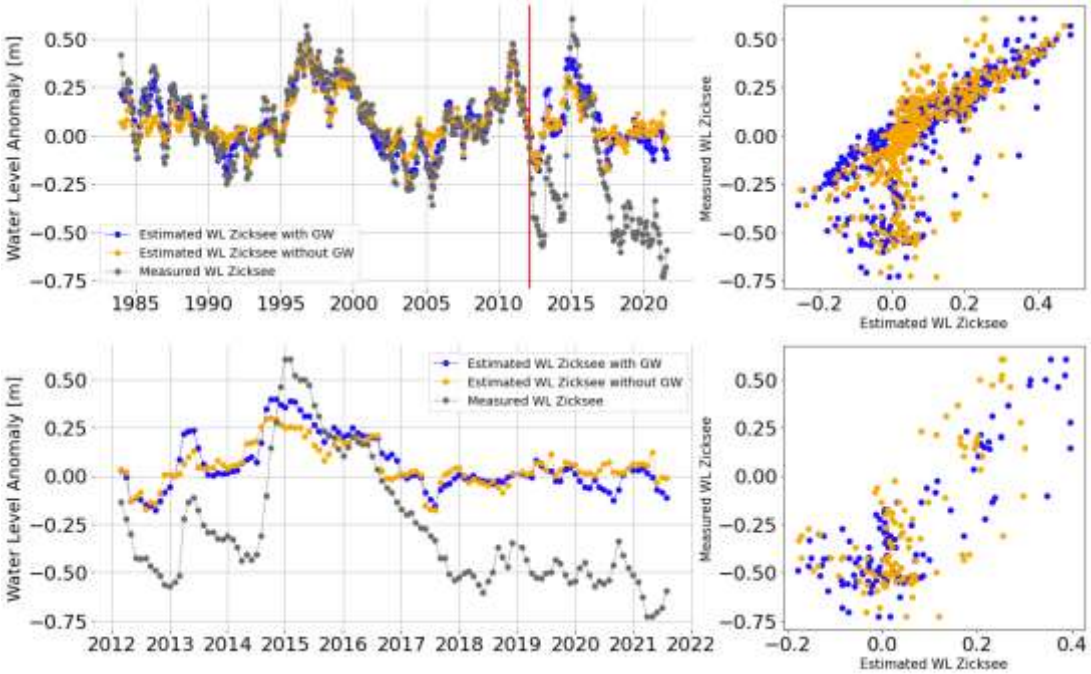


Abb. G-15: Regression result for *Zicksee*. The red line shows the separation between training (before) and test period (after).

Although performing better than *Zicksee* during the test period, the model for *Darscholacke* is has similar problems around the year 2020 (Abb. G-16:). This can also be spotted in the scatterplot and the

respective correlation coefficient for the test period (0.77 and 0.49). An analogical dynamic to 2020, occurring in the training-period, was failed to properly be addressed. Even though the yearly dynamics are modelled well, the sub-yearly fraction is subject to partly large deviations and reversed dynamics. Still, the overall correlation coefficient is high (0.9 and 0.81).

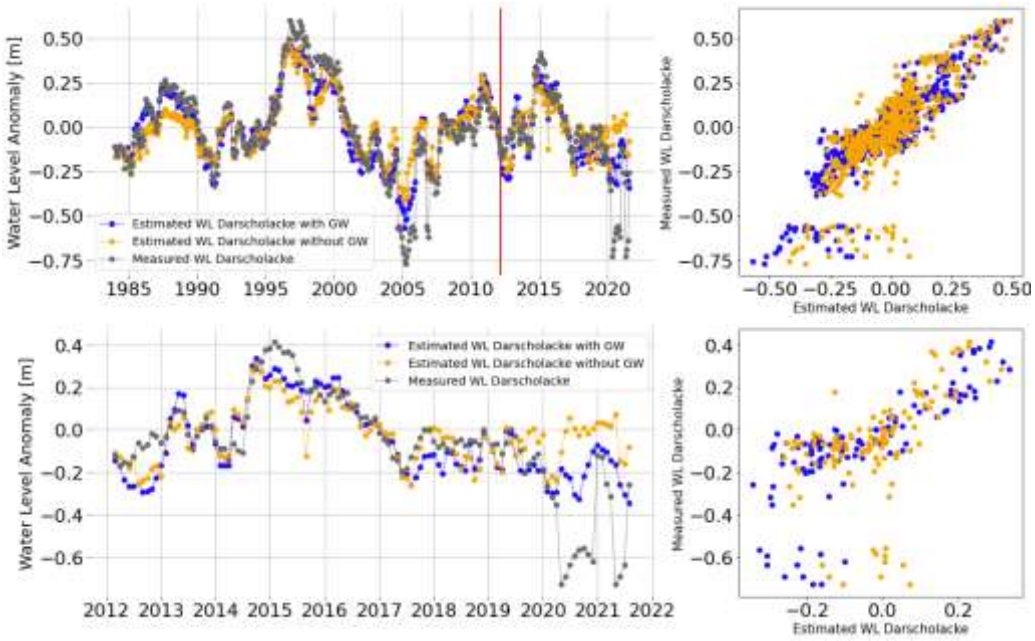


Abb. G-16: Regression result for *Darscholacke*. The red line shows the separation between training (before) and test period (after).

Although a little less pronounced and despite including more than 400 data points, a similar model instability compared to the classification can be observed. Still, many train-test splits performed in a comparable way with regard to the main model. Folds around the year 2000 (~month 190) and the year 2010 (~month 300) were particularly susceptible for changes in performance. Despite producing several outliers, the average performance settled down at around 0.1 m for all three lakes, as shown in the boxplots. Therefore, it can be concluded that different data splits have a significant influence on model performance. LOO-CV does not always follow a temporal order and the model still includes trend information which suggests alternate reasons for error susceptibility and model instability.

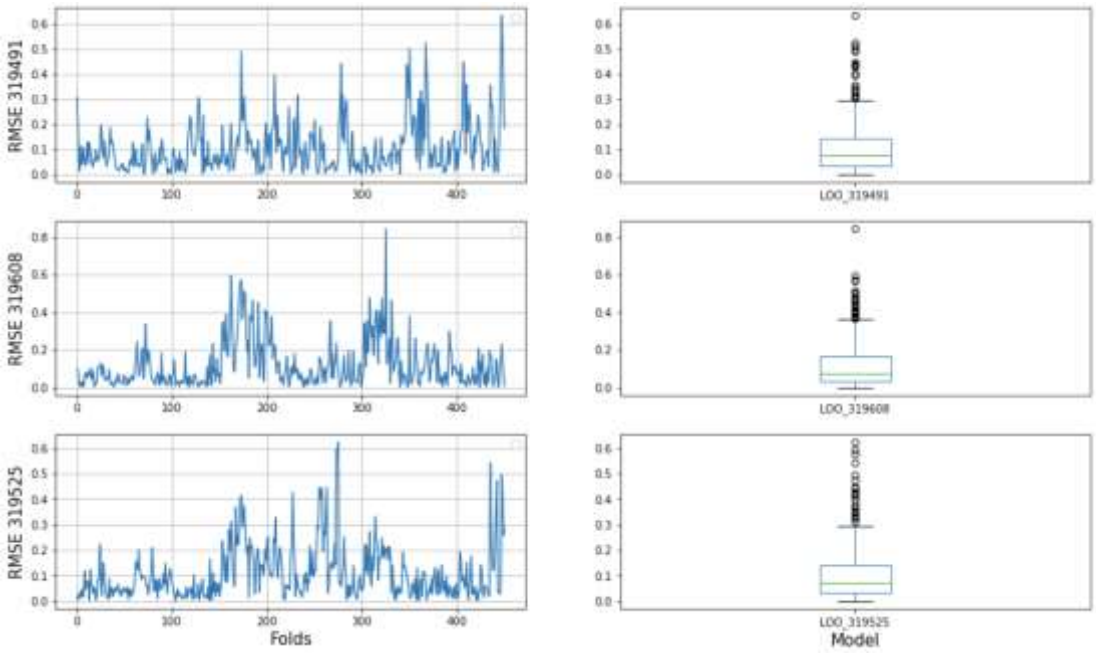


Abb. G-17: Cross-validation for all three models.

As discussed in G-4.3.2, the regression tasks are intended to enable interpretability. As these observations are based on performance evaluation and not the variability of targets when changing the input, these results are subject to careful interpretation. By plotting the feature importances (Abb. G-18:, Abb. G-19:), the dependency on information on groundwater becomes apparent: the four most important predictors are based on direct information on groundwater, whereas the SPI 6 and SPI 24 inform the model in a similar way. All other features lack importance, only information on temperature and evaporation exhibit elevate themselves to some extent. The results suggest a strong connection between WL and groundwater that meets the expectations expressed in G-5.2. Precipitation is showing its impact on a very small scale (as can also be seen in the Partial Dependency Plot (PDP)). When removing direct information on groundwater altogether, SPI 6 and especially the SPI 24 become more important and likely act as proxies for the missing groundwater data, while the other variables remain dispensable.

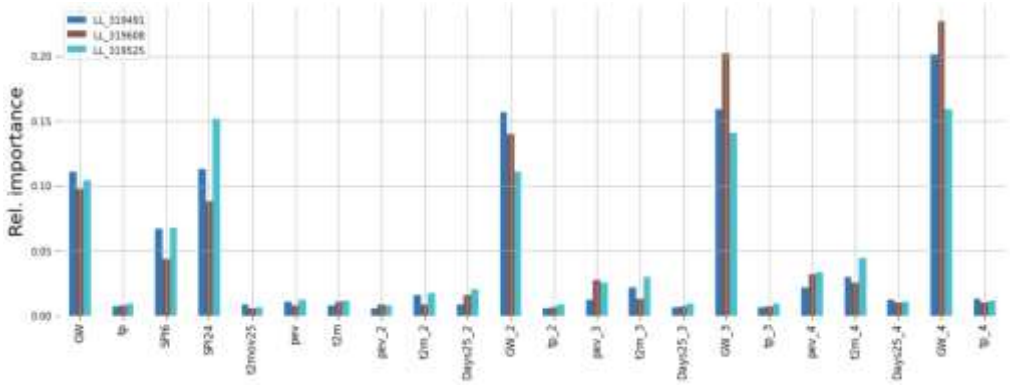


Abb. G-18: Feature Importance for all three salt pans (WL) with groundwater-contribution.

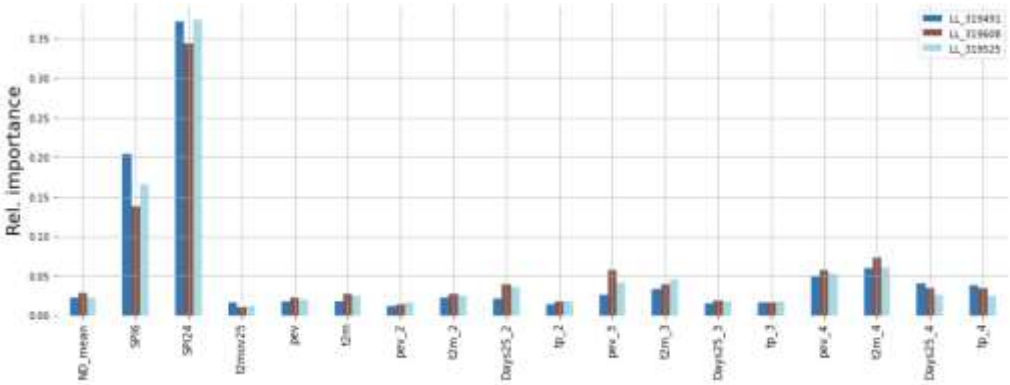


Abb. G-19: Feature Importance for all three salt pans (WL) without groundwater-contribution.

The PDPs (Abb. G-20:) suggest that SPI 24 is impacting the model in a similar fashion as groundwater integrated over, e.g., 4 months. This corresponds to the expectation gained from the literature research that groundwater inflow only happens, when a certain height (groundwater level anomaly of -0.8 m) compared to the impermeable layer is attained.

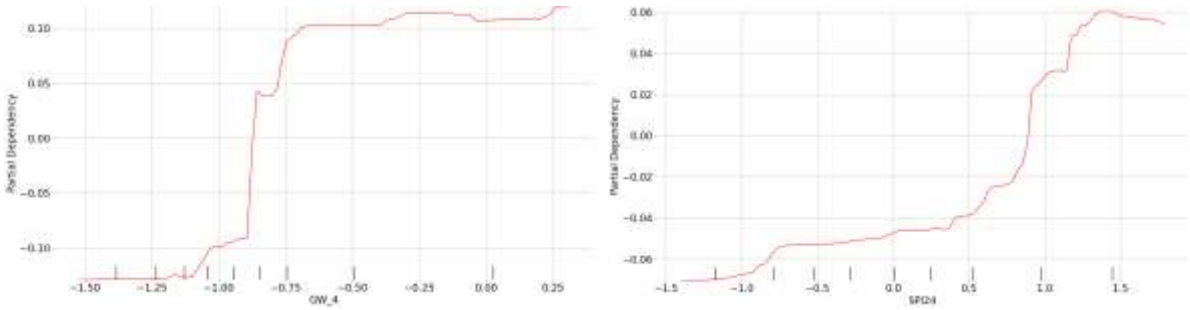


Abb. G-20: Partial Dependency Plots (PDPs) for groundwater (GW_4) and drought index SPI24 for *Lange Lacke*.

G-5.3.2 *Modelling Lake Surface Area (LSA)*

G-5.3.2.1 *Modelling LSA via a Classification Approach*

The classification results include the classification performance averaged across all wetlands. In a general sense, the model performs favourable for all eight models in respect to the test set (and more so in respect to the training set (Tab. G-9:). The mean sensitivity (meaning the true negative rate) for the groundwater model is high at around 0.78. The capability of predicting the dry-up of wetlands increases when adding information on later months, starting with a value of 0.75 for March and finishing at 0.84 for June. A similar pattern can be observed for the model without groundwater, although with slightly lower sensitivity values. The positive predictive value (ppr) is around 0.78 proving the models balanced performance due to conditioning with reference classes as well as results classes. Focusing on the groundwater-based models, although the June model performs best, merely a small performance difference compared to the March model can be observed. Moreover, the model is improving evenly with the introduction of additional months. For June, with groundwater included, the F1-score implies an 81 % probability of estimating the correct class affiliation. In comparison, when removing the groundwater information from the model, the probability decreases by 3%. This, again, goes in accordance with the expectation, as we are merely using a proxy, the SPI 24, for groundwater. Furthermore, the models perform differently for different wetlands across all modelling experiments. Wetlands with a high variability in drying and replenishing have a higher impact on overall model performance, as stable wetlands are easier for the model to predict (similar effect as observed by Daniel et al. (2022)).

Tab. G-9: Performance for eight different classification models.

Model	March	April	May	June
F1 Testscore with GW	0.76	0.76	0.78	0.81
F1 Trainscore with GW	0.91	0.92	0.92	0.92
F1 Testscore w/o GW	0.74	0.74	0.74	0.78
F1 Trainscore w/o GW	0.9	0.92	0.91	0.92

Three categories of wetlands emerge (based on model for June): one with satisfying results (< two misclassifications during the testing period), one with low model accuracy (> two misclassifications) and one composed of merely a single class (meaning that the salt pans either has not fallen try during the observed period or is always dry). These include: *Huldenlacke, Gsigsee, Darscholacke, Kuhbrunnlacke, Unterer Stinkersee, Kirchsee, Kleine Neubruchlacke, Hottergrube, Badesee Apetlon, St. Martins Therme 2*. The first group consists of the following salt pans: *Lange Lacke, Neubruchlacke, Kiesgrube Standlacke, Wörthenlacke 2, Heidlacke, Sechsmahdlacke, Martinhoflacke, Fuchslochlacke 1 and 2, Herrnsee, unbekannt, Zicksee* und *Zicklacke*. Part of the second group are: *Ochsenbrunnlacke, Wörtenlacke 1, Mittlerer Stinkersee, Oberer Stinkersee, Fuchslochlacke 3, Hochstätten, Birnbaumlacke, Kaschitzlacke* und *Albersee*. These results can be compared to the ecological classification from (Naturschutzbund Burgenland, 2012) that hints at anthropogenic forcing, although no firm conclusion could be drawn here. This line of thought is being followed more extensively in the regression task on LSA. The individual results can be seen in Abb. G-21:.

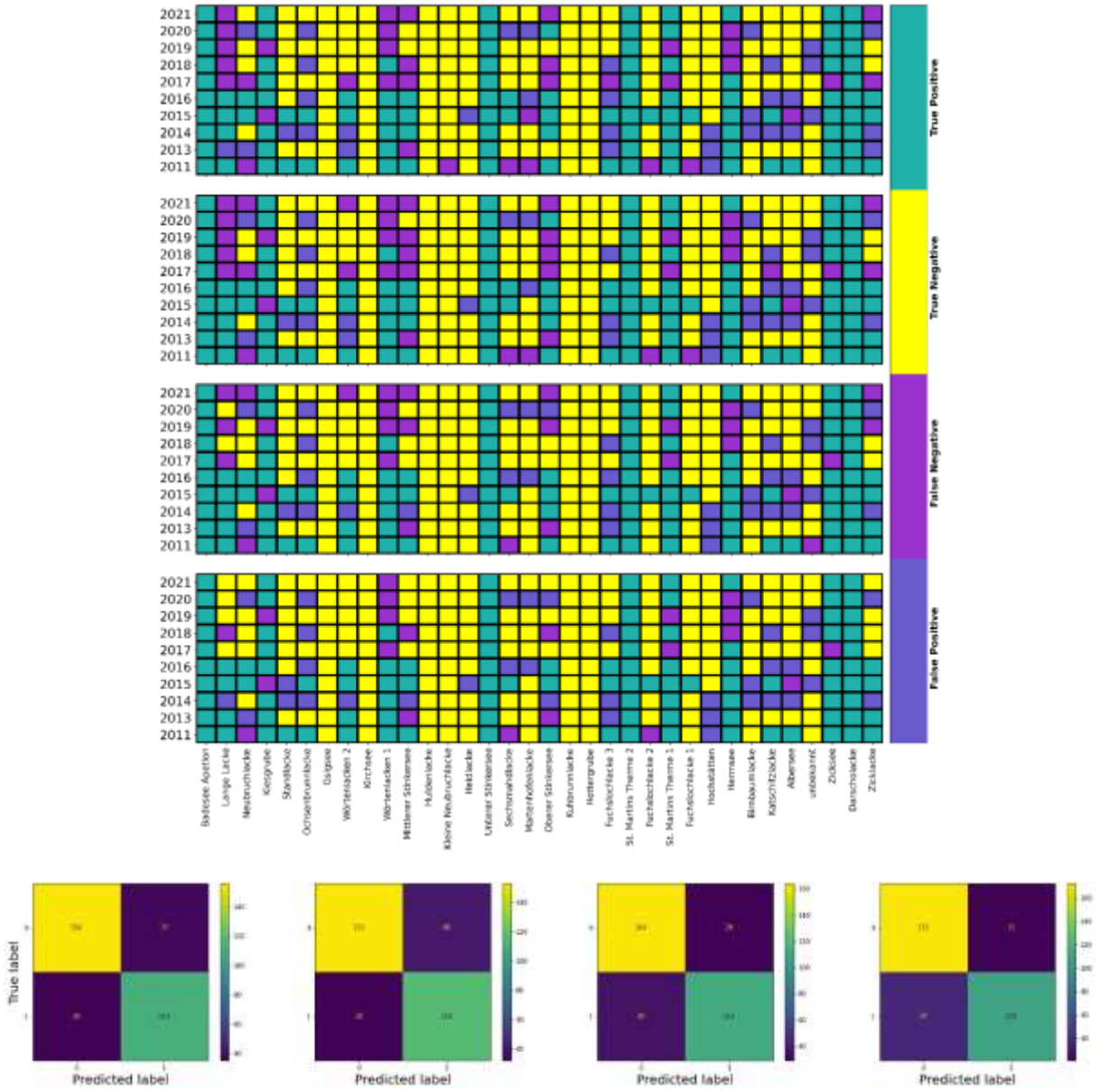


Abb. G-21: Top figure: Classification results with Groundwater included for the models March(top), April, May and June(bottom). The colours show which wetland was correctly classified in which year. Bottom figure: respective confusion matrices for the month March (left) to June(right). The results correspond to the null hypothesis stated in G-4.3.1.1.

Rather minor changes develop when excluding information on groundwater (Abb. G-22:). For some lakes, the results stay the same, while for others, a single-year prediction, mostly impairing the performance, changed. Additional misclassifications for the model in June appeared for *Lange Lacke, Neubruchlacke, Standlacke, Wörtenlacke 2, Mittlerer Stinkersee, Sechsmahdlacke, Martinhoflacke, Fuchslochlacke 3, Kaschitzlacke, St. Martins Therme 1* and *Zicklacke*. For some, like *Herrnsee*, the results improved by one additional correct classification.

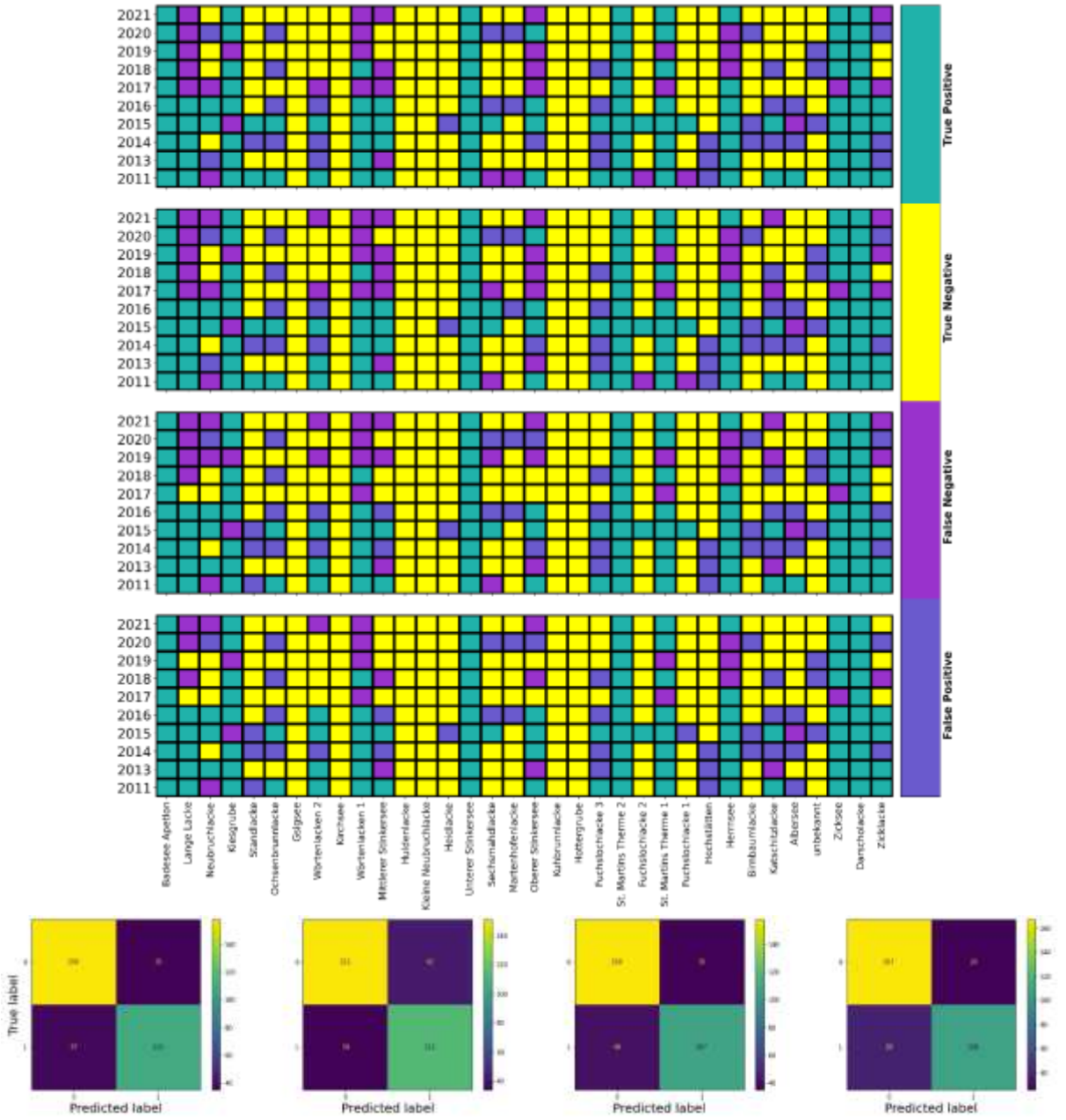


Abb. G-22: Top: classification results without groundwater included for the models March (top), April, May and June(bottom). The colours show which wetland was correctly classified in which year. Bottom: respective confusion matrices for the month March (left) to June (right). The results correspond to the null hypothesis stated in G-4.3.1.1.

G-5.3.2.2 *Modelling Lake Surface Area (LSA) via a Regression Approach*

The regression model targeting LSA is, in many ways, closely related to the model targeting WL. Most importantly, 34 instead of three salt pans were subject to modelling and subsequent interpretation. During the testing period the model performances show ample variability between the observed salt pans (Abb. G-23:). Hence, the trained models do not inherit the ability to generalize between the salt pans. Evaluating the test-performance for both setups, the RMSE relative to the respective mean salt pan area averages in at 0.99, and 0.93 respectively. *Lange Lacke* achieves a relative RMSE of 0.5, *the Unterer Stinkersee* 0.4, the *Zicksee* 0.22 and the *Darscholacke* 0.28. The omittance of groundwater-related variables worsens the results for nearly every salt pan, although not in a major way.

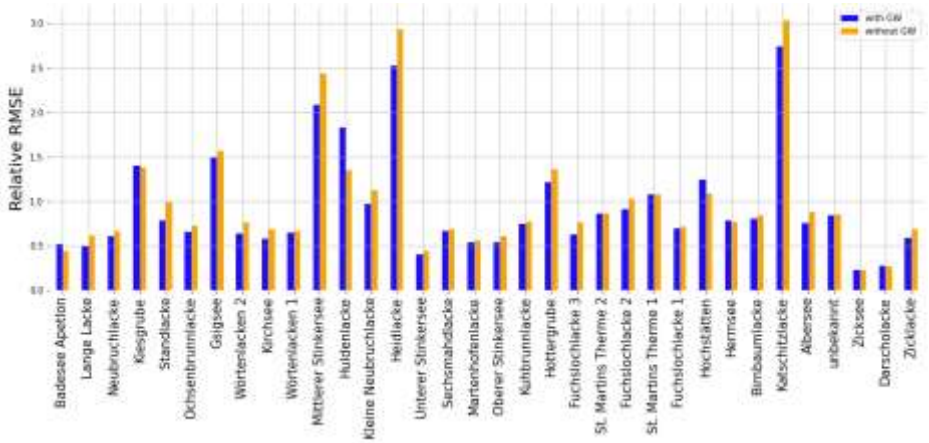


Abb. G-23: RMSE for the test period relative to salt-pan wise average LSA over the whole period for the model with groundwater contribution (blue) and without groundwater contribution (orange)

Comparing the RMSE between train and test periods, most models show a similar magnitude of error, although some perform a great deal worse than others for the test set (Abb. G-24:). These include *Kiesgrube, Mittlerer Stinkersee, Huldenlacke, Katschitzlacke* and *Heidlacke*. This difference in magnitude can be explained when studying the individual estimations in more detail: in most cases, a high RMSE was connected to misclassifications and/or the particularly wet year 2015, both resulting in outliers that were not estimated correctly. Furthermore, for some smaller lakes, the classifier merged two lakes, as in the case of *Lange Lacke* and *Katschitzlacke* in April, May and June 2015 resulting in extraordinarily high RMSE. Averaging the error over all individuals, the model tested over the train period achieves a relative RMSE of 0.49, and 0.57 for the respective model without groundwater level.

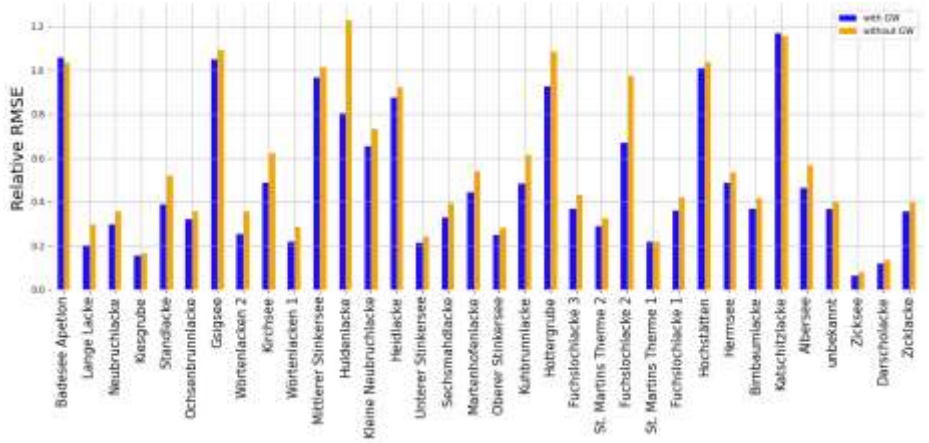


Abb. G-24: RMSE for the training-period relative to salt-pan wise average LSA over the whole period.

Plotting four exemplary salt pans provides further insight (Abb. G-25:). *Lange Lacke* exhibits a high natural variability that is depicted well in the training dataset. Still, the regression fails to follow the decrease in LSA from 2018 onwards. The dynamics of the three other wetlands, whose water extent dynamics are less pronounced, can be captured in a less erroneous manner, although outliers often could not be reproduced.

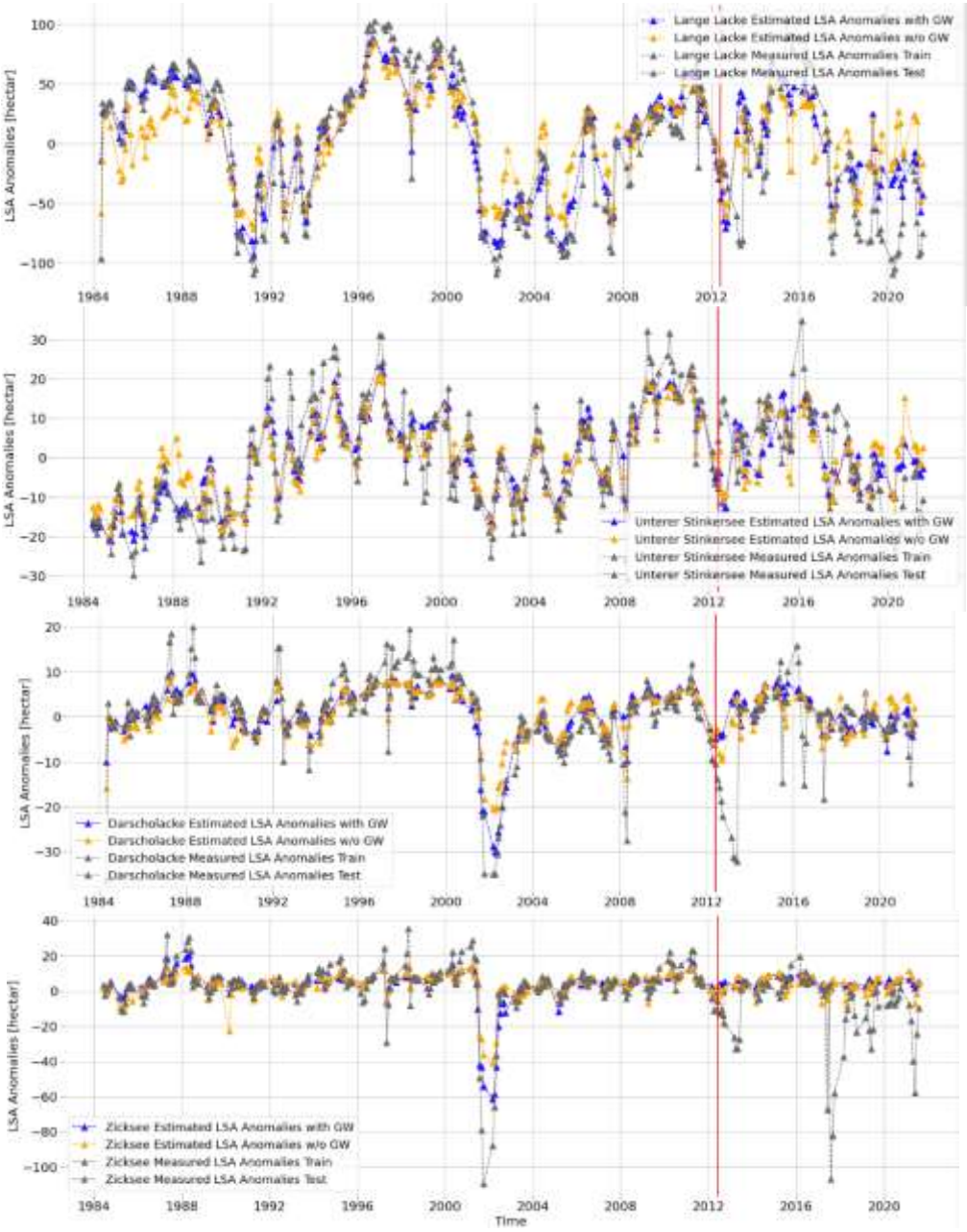


Abb. G-25: Regression results for *Lange Lacke, Unterer Stinkersee, Zicksee* and *Darscholacke*. The red line shows the separation between training (before) and test period (after).

Correlation analysis is revealing similar insights compared to G-5.3.1, only in an exacerbated manner (see Tab. G-10:). The average correlation over all salt pans is much smaller for the test-set than for the training set. As hyperparameter-tuning was employed to prevent overfitting, other effects are likely to have accumulated during modelling. Firstly, the remotely sensed time series come with certain inaccuracies and noise hindering the accurate modelling. Secondly, the interpolation of missing months is likely to not constitute a realistic representation of the processes on the ground. Thirdly, the condition of the micro-ecosystems might have changed between the test- and train-period, resulting in a different response to the applied predictors. As a general remark, performance for artificially fed salt pan was better than for more natural-fed ones. Furthermore, small salt pans tend to score worse than larger ones. Relating the hydro-ecological status given by Naturschutzbund Burgenland (2012) can, as an example, be done for *Oberer-*, *Mittlerer* and *Unterer Stinkersee*. The former and latter one received grades 2 and 3, therefore displaying a tendencially healthy ecological condition, whereas *Mittlerer Stinkersee* is graded a 4. Interestingly, the models for the salt pans in a better ecological shape tend to perform better, as can be seen by the error-bars in Abb. G-23: and Abb. G-24:, although *Mittlerer Stinkersee* has a smaller size compared to the other two.

Tab. G-10: Pearson's r correlation coefficient between measured water level and estimated water level for all six model setups.

Corr. coeff	Lange Lacke GW	L. Lacke	Unt. Stinker - see GW	Unt. Stinker -see	Dar.- lacke GW	Dar.- lacke	Zicks. GW	Zicks.	All salt pans GW	All salt pans
WL - Overall	0.89	0.8	0.85	0.82	0.77	0.77	0.67	0.66	0.83	0.8
WL - Train	0.95	0.9	0.93	0.92	0.91	0.9	0.91	0.89	0.93	0.91
WL - Test	0.68	0.46	0.45	0.35	0.23	0.27	0.32	0.19	0.48	0.34

Feature importances are again showing the importance of groundwater levl, may it be via *in-situ* measurements or by the way of SPI 24 as a proxy (Abb. G-26: and Abb. G-27:). This holds true for all four exemplary salt pans, although some variation is visible. E.g., the original groundwater time series is receives a higher weight for *Lange Lacke* and *Unterer Stinkersee*, whereas averaging it over 4 months turns out more helpful for *Lange Lacke* and *Darscholacke*. This is likely due to the lower variability in the observed LSA time series of the latter two wetlands. A more detailed analysis shows that groundwater averaged over 4 months is most important among all variables, followed by the SPI 24 and the original groundwater time series. In parallel to the results in G-5.3.1, the FI without groundwater information emphasizes the importance of the SPI 24 as a replacement. The interest tends to be more equally distributed in the latter case, as more salt pans exhibit higher values for former peripheral features.

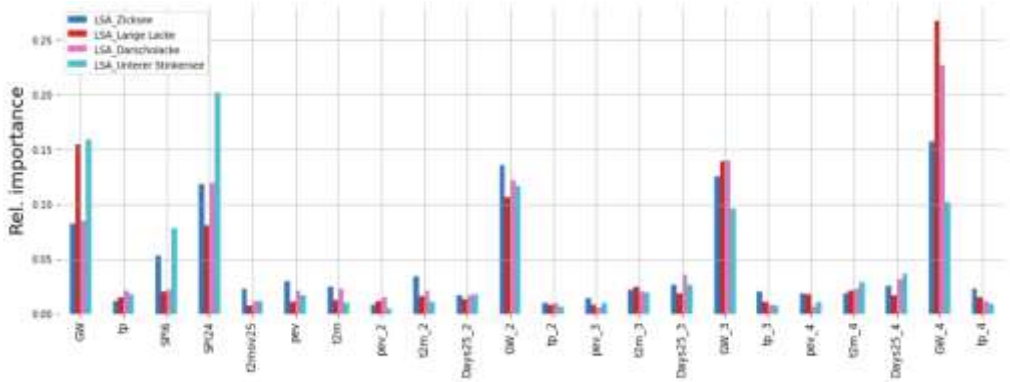


Abb. G-26: Feature Importance (FI) for four exemplary salt pans with included information on groundwater

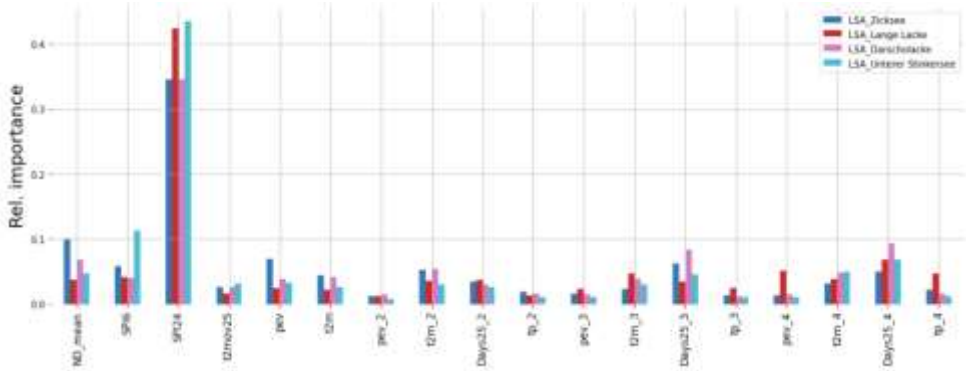


Abb. G-27: Feature Importance (FI) for four exemplary salt pans with excluded information on groundwater.

G-6 Dissemination of project results

G-6.1 Presentations at scientific conferences

The project idea, methods and preliminary results have been presented at the following scientific conferences:

- Schlaffer, S., Schauer, H., Büechi, E., & Dorigo, W. (2022). Fernerkundungsbasiertes Langzeitmonitoring der Wasserflächen im Nationalpark Neusiedler See - Seewinkel. Tagungsband 22. Klimatag, 90–91.

- Schauer, H., Schlaffer, S., Büechi, E., and Dorigo, W. (2022): Remote sensing-based monitoring of the water surfaces in the Neusiedler See – Seewinkel National Park, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-11157, https://doi.org/10.5194/egusphere-egu22-11157.

G-6.2 Dissemination activities with stakeholders

On 11 May 2022, three project members had a meeting with the administration of the Nationalpark (NP) Neusiedler See - Seewinkel. The project activities and some preliminary results were presented and discussed. The representatives of the NP administration expressed great interest in the activities, such as the automatic delineation of water-covered area per Salzlacke and the forecasting of a complete drying of wetlands during summer. The project members were provided with information on additional monitoring programmes being carried out by other stakeholders and feedback was given concerning the uncertainties of in-situ water level measurements, especially in cases where the water level gauge is not installed at the deepest point of a wetland. Three different wetlands were visited *(Illmitzer Zicksee, Unterer Stinkersee, Lange Lacke*) and an overview of the main ecologic and hydrologic problems was given for each one by the NP representatives.

G-7 Conclusions

The FEMOWinkel project has produced results which hold important implications for future wetland monitoring and management in the *Seewinkel*. Monitoring of wetland water surface area using Landsat data is feasible with approximately monthly intervals. Limitations of the approach relate to the incomplete masking of clouds in the Landsat surface reflectance products which were used in this project. While the focus in FEMOWinkel was to provide a long time series reaching back as far as the mid-1980s, for future operational monitoring activities data from the European Sentinel-2 mission are available since 2016 offering higher spatial and temporal resolutions. Especially the latter will help to overcome limitations related to cloud cover.

The data-driven modelling approach has proven to be a performant alternative to simple statistical and physical models for this case study, although the performance is heavily reliant on data pre-processing and model setup. Although the individual, monthly salt pan dynamics have not been accurately represented by performing Random Forest (RF) regression analysis targeting the WL and LSA, the RF classifier, as a baseline, was able to capture the inter-annual dynamics. Hence, a robust prediction whether a salt pan will or will not fall dry during summer with models with input from the months March to June seems possible, regardless of whether groundwater level is used as a model feature. Problems concerning overfitting and the change of environmental conditions of individual salt pans between training and testing period were identified as particularly important for the classification task. Additional challenges particularly apply to the regression experiments, e.g., data quality, model timing, the difference in performance between smaller and larger salt pans and issues concerning the modelling algorithm itself, namely its tendency for overfitting. Intermediate model setups, e.g. the seasonal prediction of drying events or the prediction of the first drying event during a year are likely to perform in correspondence to its degree of generalization. However, the drying trend of recent years is not consistently captured in all the models. Especially models that do not incorporate information on groundwater have problems in this respect. Anthropogenic water use and drainage could not be incorporated into the models due to a lack of available data. Such information would be especially useful in cases where groundwater level is not solely dependent on the climatic water balance but more and more influenced by anthropogenic factors, such as growing water use for irrigated agriculture. Data-driven modelling could incorporate such additional data in a more or less direct manner. We see an increased availability of open data on anthropogenic water use as one of the main bottlenecks for further use of such models for making robust forecasts on the future eco-hydrological state of the *Salzlacken*.

G-8 References

Albarqouni, M. M. Y., Yagmur, N., Balcik, F. B., & Sekertekin, A. (2022). Assessment of Spatio-Temporal Changes in Water Surface Extents and Lake Surface Temperatures Using Google Earth Engine for Lakes Region, Turkiye. ISPRS INTERNATIONAL JOURNAL OF GEO-INFORMATION, 11(7). https://doi.org/10.3390/ijgi11070407

Albert, R. . (2013). Die salzliebende (halophile) Vegetation des Seewinkels. In J. . Fally & L. Karpati (Eds.), NationalparkNeusiedlerSee–Seewinkel.Fertő–Hanság Nemzeti Park. Monographische Studien über das Gebiet Neusiedler See und Hanság (pp. 122–131).

Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization. J. Mach. Learn. Res., 13(null), 281–305.

Birkett, C. M. (2000). Synergistic Remote Sensing of Lake Chad Variability of Basin Inundation. Remote Sensing of Environment, 72(2), 218–236. https://doi.org/10.1016/S0034-4257(99)00105-4

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

BUNDESMINISTERIUM FÜR NACHHALTIGKEIT UND TOURISMUS. (2018). EHYD MESSSTELLEN UND ARCHIVDATEN DER HYDROGRAPHIE ÖSTERREICHS.

Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: an empirical study on performance estimation methods. Machine Learning, 109(11). https://doi.org/10.1007/s10994-020-05910-7

Chang, F.-J., Chen, P.-A., Lu, Y.-R., Huang, E., & Chang, K.-Y. (2014). Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. Journal of Hydrology, 517, 836–846. https://doi.org/https://doi.org/10.1016/j.jhydrol.2014.06.013

Cheng, F. Y., & Basu, N. B. (2017). Biogeochemical hotspots: Role of small water bodies in landscape nutrient processing. Water Resources Research, 53(6), 5038–5056. https://doi.org/10.1002/2016WR020102

Chivas, A. R. . (2007). Chapter 10 Terrestrial Evaporites. In S. J. Nash, D.J.; McLaren (Ed.), Geochemical Sediments and Landscapes (pp. 330–364). Blackwell Pub.

Choi, C., Kim, J., Han, H., Han, D., & Kim, H. S. (2020). Development of Water Level Prediction Models Using Machine Learning in Wetlands: A Case Study of Upo Wetland in South Korea. Water, 12(1), 93. https://doi.org/10.3390/w12010093 M4 - Citavi

Convention on Wetlands. (2021). Global Wetland Outlook: Special Edition 2021. https://www.global-wetland-outlook.ramsar.org/s/Ramsar-GWO_Special-Edition-2021ENGLISH_WEB.pdf

Daniel, J., Rooney, R. C., & Robinson, D. T. (2022). Climate, land cover and topography: essential ingredients in predicting wetland permanence. Biogeosciences, 19(5), 1547–1570. https://doi.org/10.5194/bg-19-1547-2022

Delaney, C., Li, X., Holmberg, K., Wilson, B., Heathcote, A., & Nieber, J. (2022). Estimating Lake Water Volume With Regression and Machine Learning Methods. Frontiers in Water, 4. https://doi.org/10.3389/frwa.2022.886964

Dick, G., Dvorak, M., Grüll, A., Kohler, B., & Rauer, G. (1994). Vogelparadies mit Zukunft? Ramsar Bericht 3. Neusiedler See - Seewinkel. UBA.

Forkel, M., Dorigo, W., Lasslop, G., Teubner, I., Chuvieco, E., & Thonicke, K. (2017). A data-driven approach to identify controls on global fire activity from satellite and climate observations (SOFIA V1). Geoscientific Model Development, 10(12), 4443–4476. https://doi.org/10.5194/gmd-10-4443-2017

Foti, R., del Jesus, M., Rinaldo, A., & Rodriguez-Iturbe, I. (2012). Hydroperiod regime controls the organization of plant species in wetlands. Proceedings of the National Academy of Sciences, 109(48), 19596–19600. https://doi.org/10.1073/pnas.1218056109

Gardner, R. C., Archiesi, S., Beltrame, C., Finlayson, C. M., Galewski, T., Harrison, I., Paganini, M., Perennou, C., Pritchard, D. E., Rosenqvist, A., & Walpole, M. (2015). State of the World's Wetlands and their Services to People: A compilation of recent analyses (No. 7; Ramsar Scientific and Technical Briefing Note). http://www.ramsar.org/sites/default/files/documents/library/bn7e_0.pdf

Ghorbani, M. A., Deo, R. C., Karimi, V., Yaseen, Z. M., & Terzi, O. (2018). Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey. Stochastic Environmental Research and Risk Assessment, 32(6), 1683–1697. https://doi.org/10.1007/s00477-017-1474-0 M4 - Citavi

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18–27. https://doi.org/10.1016/J.RSE.2017.06.031

Goutte, C., & Gaussier, E. (2005). A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. Lecture Notes in Computer Science, 3408. https://doi.org/10.1007/978-3-540-31865-1_25

Grömping, U. (2015). Variable importance in regression models. WIREs Computational Statistics, 7(2), 137–152. https://doi.org/https://doi.org/10.1002/wics.1346

Hess, L. L., Melack, J. M., Novo, E. M. L. M., Barbosa, C. C. F., & Gastil, M. (2003). Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. Remote Sensing of Environment, 87(4), 404–428. https://doi.org/10.1016/j.rse.2003.04.001

Horváth, Z., Ptacnik, R., Vad, C. F., & Chase, J. M. (2019). Habitat loss over six decades accelerates regional and local biodiversity loss via changing landscape connectance. Ecology Letters, 22(6), 1019–1027. https://doi.org/10.1111/ELE.13260

Hrnjica, B., & Bonacci, O. (2019). Lake Level Prediction using Feed Forward and Recurrent Neural Networks. Water Resources Management, 33(7), 2471–2484. https://doi.org/10.1007/s11269-019-02255-2 M4 - Citavi

Jensen, J. R. (2007). Remote Sensing of the Environment (2nd ed.). Pearson Education Inc.

Jones, J. (2019). Improved Automated Detection of Subpixel-Scale Inundation—Revised Dynamic Surface Water Extent (DSWE) Partial Surface Water Tests. Remote Sensing, 11(4), 374. https://doi.org/10.3390/rs11040374

Junk, W. J., An, S., Finlayson, C. M., Gopal, B., Květ, J., Mitchell, S. A., Mitsch, W. J., & Robarts, R. D. (2012). Current state of knowledge regarding the world's wetlands and their future under global climate change: a synthesis. Aquatic Sciences, 75(1), 151–167. https://doi.org/10.1007/s00027-012-0278-z

Kamal, Neel; Pachauri, S. (2018). Mann-Kendall Test - A Novel Approach for Statistical Trend Analysis. International Journal of Computer Trends and Technology. https://doi.org/10.14445/22312803/ijctt-v63p104

Klemas, V. (2013). Remote sensing of emergent and submerged wetlands: an overview. International Journal of Remote Sensing, 34(18), 6286–6320. https://doi.org/10.1080/01431161.2013.800656

Krachler, R., Krachler, R., Milleret, E., & Wesner, W. (2000). LIMNOCHEMISCHE UNTERSUCHUNGEN ZUR AKTUELLEN SITUATION DER SALZLACKEN IM BURGENLÄNDISCHEN SEEWINKEL. Burgenländische Heimatblätter, 62, 1–48.

Kumar, R., Musuuza, J. L., Van Loon, A. F., Teuling, A. J., Barthel, R., Ten Broek, J., Mai, J., Samaniego, L., & Attinger, S. (2016). Multiscale evaluation of the Standardized Precipitation Index as a groundwater drought indicator. Hydrology and Earth System Sciences, 20(3). https://doi.org/10.5194/hess-20-1117-2016

Li, B., Yang, G., Wan, R., Dai, X., & Zhang, Y. (n.d.). Comparison of random forests and other statistical methods for the prediction of lake water level: a case study of the Poyang Lake in China. Hydrology Research, 47(S1), 69–83. https://doi.org/10.2166/nh.2016.264

Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2018). Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. Journal of Machine Learning Research, 18(185), 1–52.

Liu, G., & Schwartz, F. W. (2011). An integrated observational and model-based analysis of the hydrologic response of prairie pothole systems to variability in climate. Water Resources Research, 47(2), 1–15. https://doi.org/10.1029/2010WR009084

Lowenstein, T. K., & Hardie, L. A. (1985). Criteria for the recognition of salt-pan evaporites. Sedimentology, 32(5). https://doi.org/10.1111/j.1365-3091.1985.tb00478.x

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). Sen2Cor for Sentinel-2. Image and Signal Processing for Remote Sensing XXIII, 1042704(October 2017), 3. https://doi.org/10.1117/12.2278218

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425–1432. https://doi.org/10.1080/01431169608948714

Mitchell, S. A. (2012). The status of wetlands, threats and the predicted effect of global climate change: the situation in Sub-Saharan Africa. Aquatic Sciences, 75(1), 95–112. https://doi.org/10.1007/s00027-012-0259-2

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J. N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. Earth System Science Data, 13(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Naturschutzbund Burgenland. (2007). Renaturierung ausgewählter Salzlacken des burgenländischen Seewinkels.

Naturschutzbund Burgenland. (2012). Salzlacken des Seewinkels.

Nhu, V.-H., Shahabi, H., Nohani, E., Shirzadi, A., Al-Ansari, N., Bahrami, S., Miraki, S., Geertsema, M., & Nguyen, H. (2020). Daily Water Level Prediction of Zrebar Lake (Iran): A Comparison between M5P, Random Forest, Random Tree and Reduced Error Pruning Trees Algorithms. ISPRS International Journal of Geo-Information, 9(8), 479. https://doi.org/10.3390/ijgi9080479 M4 - Citavi

Niculescu, S., Boissonnat, J.-B., Lardeux, C., Roberts, D., Hanganu, J., Billey, A., Constantinescu, A., & Doroftei, M. (2020). Synergy of High-Resolution Radar and Optical Images Satellite for Identification and Mapping of Wetland Macrophytes on the Danube Delta. Remote Sensing, 12(14), 2188. https://doi.org/10.3390/RS12142188

Ogilvie, A., Belaud, G., Massuel, S., Mulligan, M., Le Goulven, P., & Calvez, R. (2018). Surface water monitoring in small water bodies: potential and limits of multi-sensor Landsat time series. Hydrology and Earth System Sciences, 22, 4349–4380. https://doi.org/10.5194/hess-22-4349-2018

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540(7633), 418–422. https://doi.org/10.1038/nature20584

Pekel, J.-F., Vancutsem, C., Bastin, L., Clerici, M., Vanbogaert, E., Bartholomé, E., & Defourny, P. (2014). A near real-time water surface detection method based on HSV transformation of MODIS multi-spectral time series data. Remote Sensing of Environment, 140, 704–716. https://doi.org/10.1016/j.rse.2013.10.008

Prigent, C., Papa, F., Aires, F., Rossow, W. B., & Matthews, E. (2007). Global inundation dynamics inferred from multiple satellite observations, 1993-2000. Journal of Geophysical Research, 112(D12), D12107. https://doi.org/10.1029/2006JD007847

R Core Team. (2017). R: A Language and Environment for Statistical Computing. https://www.r-project.org/

Reschke, J., Bartsch, A., Schlaffer, S., & Schepaschenko, D. (2012). Capability of C-Band SAR for Operational Wetland Monitoring at High Latitudes. Remote Sensing, 4(12), 2923–2943. https://doi.org/10.3390/rs4102923

Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. SN Applied Sciences, 3(2), 272. https://doi.org/10.1007/s42452-021-04148-9

Safaee, S., & Wang, J. (2020). Towards global mapping of salt pans and salt playas using Landsat imagery: a case study of western United States. International Journal of Remote Sensing, 41(22), 8692–8715. https://doi.org/10.1080/01431161.2020.1781285

Santiago Beguería, & Vicente-Serrano, S. M. (2017). Package “SPEI.”

Schwatke, C., Dettmering, D., Bosch, W., & Seitz, F. (2015). DAHITI – an innovative approach for estimating water level time series over inland waters using multi-mission satellite altimetry. Hydrology and Earth System Sciences, 19(10), 4345–4364. https://doi.org/10.5194/hess-19-4345-2015

Sharma, L. K., Naik, R., & Pandey, P. C. (2021). A Focus on Reaggregation of Playa Wetlandscapes in the Face of Global Ecological Disconnectivity. In Advances in Remote Sensing for Natural Resource Monitoring. https://doi.org/10.1002/9781119616016.ch18

Shaw, P. A., & Bryant, R. G. (2011). Pans, Playas and Salt Lakes. In Arid Zone Geomorphology: Process, Form and Change in Drylands. https://doi.org/10.1002/9780470710777.ch15

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms. arXiv. https://doi.org/10.48550/ARXIV.1206.2944

Soja, G., Züger, J., Knoflacher, M., Kinner, P., & Soja, A.-M. (2013). Climate impacts on water balance of a shallow steppe lake in Eastern Austria (Lake Neusiedl). Journal of Hydrology, 480, 115–124. https://doi.org/10.1016/j.jhydrol.2012.12.013

Soltani, K., Azari, A., Zeynoddin, M., Amiri, A., Ebtehaj, I., Ouarda, T. B. M. J., Gharabaghi, B., & Bonakdari, H. (2021). Lake Surface Area Forecasting Using Integrated Satellite-SARIMA-Long-Short-Term Memory Model. https://doi.org/10.21203/rs.3.rs-631247/v1

Sumiya, E., Dorjsuren, B., Yan, D., Dorligjav, S., Wang, H., Enkhbold, A., Weng, B., Qin, T., Wang, K., Gerelmaa, T., Dambaravjaa, O., Bi, W., Yang, Y., Ganbold, B., Gedefaw, M., Abiyu, A., & Girma, A. (2020). Changes in Water Surface Area of the Lake in the Steppe Region of Mongolia: A Case Study of Ugii Nuur Lake, Central Mongolia. Water, 12(5), 1470. https://doi.org/10.3390/w12051470 M4 - Citavi

Vanderhoof, M. K., Alexander, L. C., & Todd, M. J. (2016). Temporal and spatial patterns of wetland extent influence variability of surface water connectivity in the Prairie Pothole Region, United States. Landscape Ecology, 31(4), 805–824. https://doi.org/10.1007/s10980-015-0290-5

Vanderhoof, M. K., & Lane, C. R. (2019). The potential role of very high-resolution imagery to characterise lake, wetland and stream systems across the Prairie Pothole Region, United States. International Journal of Remote Sensing, 40(15), 5768–5798. https://doi.org/10.1080/01431161.2019.1582112

Vermote, E., Roger, J. C., Franch, B., & Skakun, S. (2018). LaSRC (Land Surface Reflectance Code): Overview, application and validation using MODIS, VIIRS, LANDSAT and Sentinel 2 data's. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 8173–8176. https://doi.org/10.1109/IGARSS.2018.8517622

Wang, Q., & Wang, S. (2020). Machine Learning-Based Water Level Prediction in Lake Erie. Water, 12(10), 2654. https://doi.org/10.3390/w12102654 M4 - Citavi

Wee, W. J., Zaini, N. B., Ahmed, A. N., & Ahmed, E.-S. (2021). A review of models for water level forecasting based on machine learning. Earth Science Informatics, 14, 1707–1728.

Wen, J., Han, P.-F., Zhou, Z., & Wang, X.-S. (2019). Lake level dynamics exploration using deep learning, artificial neural network, and multiple linear regression techniques. Environmental Earth Sciences, 78(6). https://doi.org/10.1007/s12665-019-8210-7 M4 - Citavi

Wieland, M., Li, Y., & Martinis, S. (2019). Multi-sensor cloud and cloud shadow segmentation with a convolutional neural network. Remote Sensing of Environment, 230, 111203. https://doi.org/10.1016/j.rse.2019.05.022

Wu, C., Wu, X., Lu, C., Sun, Q., He, X., Yan, L., & Qin, T. (2022). Characteristics and driving factors of lake level variations by climatic factors and groundwater level. Journal of Hydrology, 608, 127654. https://doi.org/10.1016/j.jhydrol.2022.127654 M4 - Citavi

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International Journal of Remote Sensing, 27(14), 3025–3033. https://doi.org/10.1080/01431160600589179

Xu, T., Longyang, Q., Tyson, C., Zeng, R., & Neilson, B. T. (2022). Hybrid Physically Based and Deep Learning Modeling of a Snow Dominated, Mountainous, Karst Watershed. Water Resources Research, 58(3). https://doi.org/10.1029/2021WR030993

Zhao, W. L., Gentine, P., Reichstein, M., Zhang, Y., Zhou, S., Wen, Y., Lin, C., Li, X., & Qiu, G. Y. (2019). Physics-Constrained Machine Learning of Evapotranspiration. Geophysical Research Letters. https://doi.org/10.1029/2019GL085291

Zhu, S., Hrnjica, B., Ptak, M., Choiński, A., & Sivakumar, B. (2020). Forecasting of water level in multiple temperate lakes using machine learning models. Journal of Hydrology, 585, 124819. https://doi.org/10.1016/j.jhydrol.2020.124819 M4 - Citavi