
Environmental Investigation Agency Data Scientist

About the Organisation

At the [Environmental Investigation Agency](#) (EIA) we investigate and campaign against environmental crime and abuse.

Our undercover investigations expose transnational wildlife crime, with a focus on elephants, pangolins, leopards and tigers, and forest crimes such as illegal logging and deforestation for cash crops like palm oil. We work to safeguard global marine ecosystems by addressing the threats posed by plastic pollution, bycatch and commercial exploitation of whales, dolphins and porpoises. Finally, we reduce the impact of climate change by campaigning to eliminate powerful refrigerant greenhouse gases, exposing related illicit trade and improving energy efficiency in the cooling sector.

Role Description and Responsibilities

This project seeks to develop a user-friendly AI tool for identifying individual tigers from their stripe patterns, to inform enforcement efforts and counter trade in tiger skins, carcasses and individuals.

EIA has an existing image database of 158 unique individual tiger skins following encounters with illegal wildlife traders in physical markets and online. Each skin is linked to one or more trader. We also collect images of seized tiger skins and carcasses. New skins are periodically added to the database and manually cross-referenced to determine if they are duplicate images, or duplicate skins but in different images; either of which can reveal dynamics of interest to law enforcement.

As the database continues to grow, manual cross-referencing is prone to error. An AI tool to manage this cross-referencing would mean that not only can we cross-reference new skins / stripe profiles that our staff obtain, but we can assist other researchers and organisations similarly gathering such images. Ultimately, this could include cross-referencing stripes of skins of dead tigers against stripes of live tigers in captivity. The EIA database of seized specimens or those offered for sale is also shared directly with governments that host databases of live wild camera-trapped images.

For the purpose of this internship, EIA will further build the database and aims to secure up to 2000 images of tiger stripes from: live animals of known individual wild tigers from amateur and professional photographers; live captive tiger images secured by EIA and partner staff and other sources; skins, carcasses and taxidermy specimens seized, from open source media; and skins, carcasses and taxidermy specimens offered for sale physically and online.

TIN – DSG

As much as possible, multiple images of the same specimen from different angles will be obtained. The images will have been annotated with bounding boxes.

The Intern will work within the EIA Campaigns Department as the sole data scientist on the project, working alongside EIA Campaigns, Intelligence and Communications personnel who have sourced and labelled the images.

The intern will use the images in the EIA database to develop an algorithm that identifies and differentiates tigers from their individual stripe patterns, and issues an alert when a potential match between a new image and a known tiger in the database is found. This could build on prior work, such as the challenge from the CVWC 2019 workshop on detecting and identifying images of Amur tigers in China, including by incorporating their public dataset of 8000 tiger images.

Expected Outcomes

At a minimum, the outcome of the Internship will be a greater understanding for all parties of the challenges in developing an accessible AI tool to scan and cross-reference stripe patterns from images of tiger skins / carcasses, against other images of skins / carcasses and against images of live tigers (at least captive ones).

In the event it is possible during the course of the internship to develop an algorithm to cross-reference stripe patterns between dead specimens, and between dead specimens and live tigers, and determine if there any matches, there are significant benefits for conservation. It opens up the opportunity for resource-constrained researchers to collaborate with each other and law enforcement; stimulating bilateral and multilateral cooperation to investigate and disrupt tiger trade.

The desired end goal is an AI tool that can process incoming images into a database, analyse the stripe patterns, and determine if an image of that particular stripe pattern profile exists in the database already. If so, it should issue a prompt to the user, noting the relevant SkinID numbers, so that the user can then extract the source information from a separate intelligence database for further analysis regarding the relationship between the sources of the matching stripe pattern profile images. The users will be EIA Campaigns, Intelligence and Investigations personnel who are not data scientists. Accordingly the AI tool will have to a user-friendly interface.

Depending on the time available, the benefits of a neutral intergovernmental body hosting a central database of images of live wild camera-trapped tigers secured by governments, with images of tigers and tiger skins in trade secured by NGOs and other researchers, could be described and the technical steps that would be required to achieve that explored and mapped out.

Last but not least, the internship will be instrumental in the preparation for a Data Study Group challenge that will follow on and expand the intern's work with the organisation. The intern will have the opportunity to represent the organization during the Data Study Group.

Supervision and Mentorship

The intern will be supervised by EIA's Campaign Leader for Tigers & Wildlife Crime. Although there are no data scientists with experience in this field at EIA, some technical support can be

TIN – DSG

offered by the Turing Institute.

EIA will ensure the Intern is exposed to a range of stakeholders at EIA and among partners to explore further how their skills can be applied to help combat wildlife crime. EIA would encourage and support the Intern to prepare papers either for academic publication, and/or, via EIA website on the challenges and final outcomes.

Person Specification

The ideal candidate is a Data Scientist with prior experience in computer vision (specifically object detection and image classification), who relishes a challenge, and is motivated by the idea of putting their skills to good use to the benefit of endangered tigers. They will not be daunted at the prospect of working with non-technical personnel, will enjoy problem-solving, are patient and able to communicate in non-technical terms. They will be comfortable with limited face-to-face time with their direct report.

Internship Logistics

The term of the internship is six months full time commencing in XX.

The salary is £30,000 p/a pro rata.

EIA operates a flexible working policy and the Intern is welcome to work from EIA's London office but note that the Campaign Leader for Tigers & Wildlife Crime works remotely and will be in the office approximately one week in every six. Communications will otherwise be via regular Microsoft Teams calls.

The images for the project will be provided in a secure FileDrop on EIA's Sharepoint, which the Intern will have password protected access to.

Global Witness Data Scientist TINDSG-002

#machinelearning #datastudygroup

About the Organisation

At [Global Witness](#) our goal is a more sustainable, just and equal planet. We want forests and biodiversity to thrive, fossil fuels to stay in the ground and corporations to prioritise the interests of people and the planet.

Role Description and Responsibilities

The goal of this internship project is to build a machine-learning classification algorithm for accurately detecting the presence of a specific type of unregulated and destructive mining of critical minerals we are investigating using high-resolution (Planet) and medium resolution satellite imagery (Sentinel and Landsat). More information on the specific type of mining will be shared during the interview.

This will be an entirely new project from a technical perspective, but will feed into an existing strand of research on critical minerals.

This classification algorithm will be applied to historic satellite imagery over large areas to detect the scale of mining in different areas and its occurrence across time.

Global Witness has attempted a related project in the past looking at a different type of mining (see this [post](#)), but this would go further than any other previous work and address sustainability issues in one of the most significant sub-sectors of the mining industry for the renewable transition.

The sources of imagery for this project will be high-resolution imagery from Planet as well as medium resolution sources from Landsat and Sentinel.

In addition to the main project, you will also support other research teams with analytical problems that are amenable to machine-learning approaches, specifically those that use publicly available satellite imagery for the detection and monitoring of palm oil and soy plantations and other forest-risk commodity crops, destructive mining and utilisation of fossil fuel infrastructure.

The intern will be part of the Data Investigations Team, a group of analysts and programmers working across the whole organisation that research data-driven stories and build lightweight applications in the service of fighting the climate crisis and exposing corporate abuses. You can see a selection of the team's works [here](#).

As part of the internship, you will work with a range of different campaign and campaign teams at the organisation, specifically with those teams looking at fossil fuel and agricultural

TIN + DSG Spring 2022

commodity production and supply chains.

Expected Outcomes

Global Witness aims to change laws through exposés, targeted communications and advocacy. The ultimate aim of a successful project would be to change national laws relating to the production of a specific set of critical minerals.

If successful, this project will provide new information on a largely opaque process of production.

Other outputs will include articles and digital communications presenting the findings from the project and new evidence they provide concerning the production of the critical minerals in question. Significant findings will also be used to in advocacy materials — such as briefings and reports for policy makers — to push for better regulation of the mining sector in question.

The internship will be instrumental in the preparation of a Data Study Group challenge that will follow on and expand the intern's work with the organisation. The intern will have the opportunity to represent the organization during the Data Study Group.

Supervision and Mentorship

The role and the project will be supervised by the Head of the Data Investigations Team. Within the Data Investigations Team, you will work alongside a team of analysts and programmers who will provide support for data wrangling tasks as well as manual classification.

Ideal Intern

The ideal intern would be:

- an excellent data scientist that is interested in applying their skills to make the world a better place;
- highly organised and self-motivated able to coordinate and drive projects forward with limited supervision;
- a good communicator who can write up and talk about the results of their research in a way that is meaningful to the public, journalists and policy makers;
- experience building machine learning algorithms for classification of satellite imagery or other remote sensing imagery;
- take a creative approach to problem solving with an appetite for applying their skills to unusual problems.

Internship Logistics

The duration of the internship is 4 months and will be for 4 days a week (Monday to Thursday) The internship will start in November 2022.

This is a remote post, but we would expect that the intern is able to join in-person meetings in London at least once a month.

The salary is £30,000 pa pro rata.

For any queries, please contact Sam Leon, Head of Data Investigations, on sleon@globalwitness.org.

The John Muir Trust Wild Places Project Researcher TINDSG-003

#machinelearning #datastudygroup

About the Organisation

[The John Muir Trust](#) is a community focused conservation charity dedicated to the experience, protection and repair of wild places across the UK. Founded in 1983, we care for some of the finest wild places in the UK, including Ben Nevis, Helvellyn, Skye and Sandwood Bay. Over 25,000 members, supporters and partners contribute to our work.

Through collaboration with local partners, the John Muir Trust works towards landscape-scale restoration of natural processes; the expansion of native woodlands, the nurturing of rare mountain plants, and on footpath repair through some of the country's wildest places.

The John Muir Trust believes that by inspiring and connecting people with wildness, this engenders a desire to protect and conserve some of the most beautiful landscapes in the world while repairing and rewilding for future generations to enjoy.

Role Description and Responsibilities

The John Muir Trust wants to understand the extent and condition of the UK's wild places and to share this data through the production and interpretation of a wild places register and wild places standard. As part of this project, we want to develop wild place criteria, which would give objective relevance to different expressions of wild places (e.g. ecological, experiential, physical, societal). We envisage the criteria and the register would be created with input and collaboration from a wide range of data sets and engagement with a wide sector of society. Once created, we hope the register and its interpretation will encourage more people to identify wild places, think more about their condition and learn what they can do to protect them.

The Trust has previously commissioned maps of relative wildness to inform its work and the identification of 42 Wild Land Areas by Scottish Natural Heritage (now NatureScot) in 2014 provides a starting point. These maps have assessed the degree of wildness by looking at the strength of four physical attributes. Whilst this method does a good job of picking out areas according to physical qualities, it has its limitations. The 42 mapped Wild Land Areas have resulted in a binary illustration of wild areas whilst in reality, wild is a continuum beyond the mapped area. The four attributes also focus largely on physical appearance of a landscape, which means wild places that feel important to people are overlooked. Similarly, small, and not necessarily remote wild places that are important for biodiversity or ecosystem services may also be missed.

TIN + DSG Spring 2022

The John Muir Trust plans to develop an exciting new approach to understanding wild places. This project will explore and develop a new scientific process for understanding the data behind wild places, that considers the broad range of physical and social criteria that makes a place wild. We have initiated a focused Wild Places Project Team who meet regularly to work on this development phase. The intern will be a central part of this new team and be key in helping to define and develop both physical (by mapping) and social (with stakeholders) criteria, through an iterative process.

The Trust has access to an array of social and physical datasets for this challenge. These include GIS maps of the UK and national datasets on wild land areas, habitat condition surveys, noise and light pollution, and public perception surveys. We have additional data on the land the Trust manages, and we can engage with our membership base of around 11,000 individuals to gather knowledge about specific locations. There will also be the opportunity for the intern to visit some of the beautiful sites managed and cared for by the Trust, to gain a better understanding of our work.

The intern will be responsible for exploring inclusive approaches for collating these data sources and using them to help the Trust classify the nature, condition, and extent of wild places. We would also like the intern to evaluate the potential of incorporating satellite imagery and computer vision techniques for this task.

The internship will be instrumental in the preparation of a Data Study Group challenge that will develop and verify the promising research avenues identified by the intern, aiming to produce an algorithm capable of classifying the wild status of the whole of the UK. The intern will have the opportunity to represent the organization during the Data Study Group.

Expected Outcomes

The successful completion of this challenge would raise the profile of the UK's wild places as a key solution to the climate and biodiversity crises and raise awareness of wildness and its benefits through ecosystem functions and services, nature-based solutions and human experience and engagement with wild places. Wild places have huge mental and physical benefits for large sections of society. Engaging with society through this work will help all recognise and value wild places for all.

From an operational perspective for the Trust, the successful completion of this project would allow us to use the dataset and register of wild places to target interventions, such as purchasing land, engaging with communities, and developing/enforcing standards & policy. With these tools we would drive up the standard of wild places benefitting wild places and society.

Supervision and Mentorship

The Intern would be directly line managed by the Research & Data Manager and part of the Wild Places Register team supported by the Director of Policy, Senior Policy Officer and other staff. The intern would be part of the wider Trust team and be invited to attend staff meetings and gatherings.

Ideal Intern

We would benefit from an intern that would provide a fresh perspective and scientific expertise with experience working with large datasets and maps.

TIN + DSG Spring 2022

The ideal candidate will have:

- Strong interpersonal skills and innovative problem solving abilities
- Ability to take initiative in their work

Any experience working with the following would also be advantageous:

GIS, maps, and spatial analysis

- Computer vision
- Satellite imagery.
- Experience combining and extracting insights from disparate datasets.

Internship Logistics

The salary is £30,000 p/a pro rata.

Location: The Intern will have the choice to work remotely or work at one of our offices in Edinburgh or Pitlochry.

Internship duration: 6 months

Start date: September 2022

Hours: Full time, but flexible working available subject to discussion

Point of contact: crstin.lambert@johnmuirtrust.org

Equipment: Intern will be provided with a new laptop and monitor for the duration of their time with the Trust.

Keep Wales Tidy and Keep Scotland Beautiful Data Scientist

TINDSG-004

#imagesformodelling #datastudygroup

About the Organisation

[Keep Wales Tidy](#) (KWT) and [Keep Scotland Beautiful](#) (KSB) are both charities who are committed to working with local communities across the UK to protect our environment now and for the future. We have a shared vision to combat climate change, tackle litter and waste, and protect and enhance the places we live, work and visit. As part of our work KSB and KWT conduct the independent elements of the statutory surveys on litter and environmental quality on behalf of every Local Authority in Wales and Scotland respectively. The models we use for these surveys, called Local Environmental Audit and Management System (LEAMS) or LMS, produce a grading system by which governments are able to ascertain whether Local Authorities are meeting their obligations to keep their streets tidy and clear of litter.

Keep Wales Tidy collect our data through the LEAMS validation surveys. LEAMS surveys provide a 'snapshot' of litter and other local environmental quality issues across Wales and allows us to track trends over time. During the survey, we measure the quality of the local environment by recording the street cleanliness grade based on how much litter and dog fouling is present, types and sources of litter present, amount of graffiti, vandalism, fly-posting, dog fouling, weeds, detritus and chewing gum staining visible. We have also collected information on the sources of litter (e.g. individual, business, domestic or construction) and have increased our survey to collect data on single use plastics, PPE litter, verge cleanliness and drinks litter in order to inform relevant developments in policy. These surveys are carried out in person by the staff at Keep Wales Tidy and the data collected on Epicollect in order to utilise GIS to map results spatially.

Along a similar vein, Keep Scotland Beautiful also follow a comparable system for the collection of litter data. Currently sites are graded against litter grades from the Code of Practice on Litter and Refuse (Scotland) 2006 as a new method of data collection is being developed (the revised 'Litter Monitoring System' (LMS). The LEAMS process uses a combination of monitoring carried out by local authorities and Keep Scotland Beautiful. Each Local Authority conducts two LEAMS audits during each financial year and Keep Scotland Beautiful carries out a third audit to provide independent verification. Each audit assesses a randomly selected 5% sample of streets and roads. All Local Authorities but Moray are included in the programme. A subjective perception rating for litter presence is also recorded for each site which takes into consideration the full visible streetscape in context, including nearby open spaces where present. This provides a robust indicator of local environmental quality standard result for each site audited. Despite apparent differences in the Scottish and Welsh requirements for litter data collection, the LEAMS and LMS models are, for the intents and purposes of litter data collection via AI, directly comparable.

TIN + DSG Spring 202

The data collected through these surveys is then used to create policy recommendations and help inform government policy in Wales and Scotland on littering, waste and environmental policy.

Role Description and Responsibilities

There is a growing interest in the use of AI in the gathering of litter data and whether or not image recognition technologies can be used to conduct LEAMS or LMS surveying at scale. In order to facilitate this investigation, Keep Wales Tidy and Keep Scotland Beautiful have agreed to start collecting a database of images of litter so that they can be utilized in a DSG challenge around the creation of a litter recognition model which could be applied to statutory cleanliness standards. It would be one of the roles of an intern to work with us to assess the first batch of collected images, evaluate their suitability for AI purposes, and make recommendations for any improvements/changes for future data collection in order to prepare these images for use at a DSG.

An intern would also be required to work alongside the policy and research teams at KWT and KSB to research and explore possible modelling approaches that could be used to classify the cleanliness of an area from its imagery. Working with staff from KWT and KSB to look at LEAMS and LMS models for the collection of litter data, they would identify any promising approaches which could then be explored in more depth at a DSG.

Additionally, given that any model would involve the capturing of images in public places, an intern would also be asked to consider the ethical and security implications of using AI in this field and make recommendations for the safe and ethical collection and use of the data in this field. This will form a part of helping KWT and KSB shape their future policy for AI, and make recommendations for how they can incorporate AI into the organisation going forward.

Both KSB and KWT have strong links with academic institutions which could support the work of the internship (Stirling University and Cardiff University respectively). This work will form part of the wider strategic aims of Welsh and Scottish Governments who we expect to have significant interest in this work.

Working on this project offers an intern an opportunity to work on a high-profile project with national significance and develop new ways of utilising AI in a field that has thus far received little attention from the field.

Expected Outcomes

At the end of the internship there would be several outputs that would form the basis of a DSG challenge. Firstly, a database of images which are useful and suitable for usage in an AI modelling context. Secondly, a written report identifying any modelling approaches that the intern has found to be helpful or potentially applicable to the litter context. These could be used to help narrow down working ideas in a DSG challenge.

Furthermore, at the end of the internship KSB and KWT will have broadened their depth of knowledge into AI and co-produced a policy document which lays out best practice in litter data collection using AI. We would also create a 5 or 10 year plan for how AI can be utilised in our work and increase the level of detail that we can collect, this would help inform policy across Wales and Scotland (including national and local government level policy).

TIN + DSG Spring 202

Supervision and Mentorship

Keep Wales Tidy will be the lead organisation for the purpose of supervision, but a team will be established from both organisations which will include data and technical support, policy expertise and litter data experts in order to ensure that there is adequate support and communication across both organisations.

Ideal Intern

The ideal intern for this role would have experience working with image datasets and preparing images for modelling purposes. Experience working with tabular data would be beneficial but not essential. They would also have a strong sense of the ethics surrounding AI and its use in the public realm. We are also looking for someone who is passionate about our work, interested in environmental policy and excited to be working on a real-world problem.

Internship Logistics

Both organisations have flexible working policies. Whilst office space is available, as environmental organisations the majority of staff work from home and online meetings are encouraged to limit travel and carbon emissions. Therefore, the candidate could be based anywhere in the UK although some travel may be required for meetings at office locations in Cardiff or Stirling. IT equipment and support can be provided.

Keep Wales Tidy (Charity Registration No: 1082058)

35 Cathedral Rd, Pontcanna, Cardiff CF11 9HB

Lead contact: Alexander Makovics: Alexander.Makovics@keepwalestidy.cymru

Keep Scotland Beautiful (Scottish Charity No: SC030332)

Glendevon House, The Castle Business Park, Stirling, FK9 4TZ

National Oceanography Centre ML Research Engineer TINDSG-005

#machinelearning #datastudygroup

About the Organisation

We are the [National Oceanography Centre](#) (NOC) - the UK's centre of excellence for oceanographic sciences. We are a national research organisation, delivering integrated marine science and technology from the coast to the deep ocean, and are one of the top five institutions of its kind in the world. The NOC has entered a truly exciting time, moving to independent & charitable status as of 1 November 2019.

We are made up of a dynamic and vibrant community of staff covering a range of specialist fields, backgrounds, and experience. A community where each employee adopts a crucial role in furthering the aspirations, advancing the frontiers of science and knowledge through our excellent scientific research, knowledge sharing and contribution to the health of the oceans, with a focus in improving the world in which we live. Our work is balanced by our strong sense of purpose, values and behaviours and an unwavering commitment to a 'one NOC' approach.

Role Description and Responsibilities

Introduction

The project will be around the study of eddies in terms of detection, tracking, and modelling of their trajectories. Depending on the progress, the project can also be expanded into underwater gliders path planning.

The study of eddies plays a vital role for oceanographers in understanding and modelling the ocean. One area of study where a more accurate model of the ocean would contribute significantly towards, is in the understanding of the global ocean Meridional Overturning Circulation (MOC). The MOC is responsible for poleward heat transport, and deep storage of heat and carbon. Climate models generally predict that a slowdown of the MOC will occur this century, with dramatic regional and global climate changes.

Understanding, detecting, and tracking of eddies contribute not only to the oceanic sciences directly, but also to the environments of NOC's robots, such as gliders and autonomous underwater vehicles (AUVs), that are actively used for these scientific studies. Oceanographers who study eddies collect data by crossing them through with gliders to measure their temperature-salinity profile and other parameters which describe the behaviour, types, and properties of the eddies. For others, depending on the type of missions, glider pilots may want to avoid getting stuck in eddies, or use eddies to their advantage by diving with the current to significantly increase the endurance of the gliders for potentially up to multi-months and year-long missions. Path-planning of marine autonomous vehicles for environmental research could be significantly improved by determining the trajectory of eddies alongside detection and tracking.

TIN + DSG Spring 2022

About the project

You, as an intern, will get to study and work with satellite and other remote sensing datasets, as well as operational, real-world robotics datasets that are extensively used in the marine science domains such as physical oceanography, marine biology, marine biogeochemistry, and climate research. There will be opportunity to get familiar with autonomous underwater vehicles, which are increasingly becoming an important line of development for critical strategies such as net zero, monitoring offshore windfarms, validation of forecast models, and climate change studies.

The use of machine learning and deep learning in eddy detection and tracking, and modelling the trajectory is a recent topic of research with high availability of open and free data, and with high potential for further contributions and novelties. The dataset(s) for the detection and tracking of eddies would include sea surface heights and/or sea level anomaly data, but other datasets are also available including global mesoscale eddy trajectory data, Argo floats data, as well as data from glider missions provided by us at National Oceanography Centre (NOC). Additionally, fusion of features generated from classical model-based techniques could be used to improve the tracking and modelling of eddies trajectories. Model validation can be done with the available historical and live or near real time data provided with these datasets.

Team and department

You will join our Command and Control (C2) team within the Marine and Autonomous Robotics Systems group (MARS). The team is currently made up of 8 people including software engineers and researchers. Our expertise ranges from the latest front-end technologies to AI algorithms, and we really value a diverse range of backgrounds and experience. The range of experience within the team allows us to be involved in a very broad portfolio of activities, from developing user-friendly UIs to designing abstract frameworks to enable the deployment of innovative AI systems to control un-crewed robots for oceanographic research.

The C2 team develops web-based systems to remotely control and interact with autonomous underwater and surface robots. Our systems control our robots that can be deployed anywhere in the world and facilitate the transfer and processing of scientific and engineering data in near real time. This data then contributes to different leading scientific programmes and international data centres, informing scientific endeavours from climate change studies to under ice exploration.

Internship project(s)

You will have access to and help improve the NOC's Machine Learning (ML) Framework which was developed to build and maintain an eco-system for sustainable and reproducible ML models. You will have easy access to live data through our open API ecosystem which allows for more effortless data exploratory work. Your contributions to the data streams can be easily integrated by feeding in additional features generated alongside the main remote sensing datasets during the research project. The project will also enhance NOC's Automated Piloting Framework (APF), which we have developed, by enabling eddies navigation algorithm for underwater gliders.

In terms of eddies detection and tracking, currently, scientists generally use probabilistic and model-based techniques. The data for a global mesoscale eddy trajectory are available widely for use with free and open access. Path planning for gliders, on the other hand, are mostly done manually. This is an exciting opportunity for you to contribute to a greater understanding of the ocean (with improved model of the ocean) and more accelerated scientific research

TIN + DSG Spring 2022

activities. These lead to the advancement of environmental oceanic studies of exploring the ocean, improve our ability in predicting the weather, the changing climate, as well as studying their impact on the economy, fishing industries, and humanity. Furthermore, a more cost-efficient scientific expeditions will also help in the Net Zero Strategy while leading explorations of innovative technologies for tackling climate change and its impact such as carbon capture technology, and development of modern strategies by studying the interconnected, chaotic, and dynamic oceans.

Expected Outcomes

Work on eddies detection, tracking, and modelling of the trajectories can be split into multiple deliverables such as the following:

1. Work with expert physical oceanographers to understand the current available datasets and methods of detection, tracking, and modelling the trajectory of eddies (i.e., probabilistic, and dynamic models).
2. Develop a baseline implementation either from widely used model-based techniques, current state-of-the-art machine learning solution for the problem, or both.
3. Study and perform data exploratory work with the available dataset(s).
4. Collaborate with scientists and engineers in developing solutions. This can be in the form of collaborative projects or running workshops.
5. Design, develop and prototype an eddies detection and tracking algorithm, providing their trajectories and other features of interest.
6. Write and publish research output to a journal or conference.
7. Publish results of features/targets of interest to the APF to enable a path planning for gliders for eddies navigation.

Work on studying and development on the detecting, tracking, and modelling of eddies would contribute to both the environmental science research, as well as directly to field research operations being done by scientists and engineers daily at NOC. The project can contribute significantly to an increased endurance of glider operations, as well as to mission and path planning by integrating the technology into our APF. A more accurate model of the ocean would also fundamentally improve our ability to pilot multiple gliders and AUVs simultaneously and run multiple missions concurrently.

The internship will be instrumental in the preparation of a Data Study Group challenge that will follow on and expand the intern's work with the organisation. The intern will have the opportunity to represent the organization during the Data Study Group.

Supervision and Mentorship

An intern working on this challenge would have the opportunity to work directly with and be advised by experts in the physical oceanography field, as well as robotics planning and control field.

There will be various opportunities for further collaborations, as the project involves both the environmental science and robotics aspects at the National Oceanography Centre.

Ideal Intern

The ideal candidate will:

- Have a good knowledge of machine learning algorithm development and research

TIN + DSG Spring 2022

- Ideally be knowledgeable in probabilistic modelling and reasoning and have experience of working with remote sensing datasets
- Having knowledge of robotic planning and dynamics is not a must but understanding of them would be beneficial
- Be passionate about learning new things, excited about working with leading researchers and collaborating in a multidisciplinary team
- Will value quality in the things that you do.

If you're interested in joining us, but don't tick every box above, we still encourage you to apply! We're building a diverse team whose skills, experiences, and background complement one another.

Internship Logistics

The internship will offer a minimum salary of £30,000 p/a pro rata.

Location: Southampton. The internship can also be done remotely with occasional site-visits for a tour around the site to meet the team and witness the development of a range of marine robotic systems, as well as for collaborative brainstorming and development sessions.

The internship start date is anticipated to be 1st September 2022, with a duration of 6 months and it will be full time, 37 hrs/week.

Contact details:

Alvaro Lorenzo Lopez: alvaro.lorenzo@noc.ac.uk

Izzat Kamarudzaman: izzat.kamarudzaman@noc.ac.uk

The NOC is an equal opportunities employer and welcomes applications from all sections of the community. There is a guaranteed interview scheme for suitable candidates with a disability and we welcome applications from ethnic minorities currently under-represented. The NOC is an Investors in People organisation and has signed up to the Athena SWAN charter principles to take action to address gender equality.

Peak District National Park Authority

Replacing human interpretation of land cover in the Peak District with a deep convolutional neural network

TINDSG-006

#deeplearning #datastudygroup

About the Organisation

[The Peak District National Park](#) was the first of Britain's 15 national parks and is enjoyed by millions of people every year, looking for both adventure and a breathing space to escape. It is one of the busiest national parks in the world due its proximity to four major cities of the UK and its 38,000 residents. We recognise the importance of our national park and our job is to speak up and care for the Peak District National Park for all to enjoy forever.

No element of the national park landscape is untouched by past or present human activity. However, new technologies, climate change, more people and changing lifestyles mean that our potential to change the environment and the appearance of the landscape is far greater now than in any previous generation.

Currently however, there is no standard way of monitoring these changes to the landscape and this makes it difficult to target our efforts. We need to be able to measure the changes that are already occurring, as well as the effect of the improvements we make.

Role Description and Responsibilities

This research involves the investigation and development of image classification using deep-learning for land cover assessment across the Peak District National Park.

Historically, the Monitoring Landscape Change in the National Parks project (Bird et al, 2000; Taylor et al, 2000) was the last complete census of the land cover in the National Parks (approximately 14,000 km²) using medium scale aerial photography (1:20,000 - 1:25,000) flown in the early 1970s by Ordnance Survey and the late 1980s by ADAS. However, this was a massive undertaking of resource, taking almost 3 years for highly trained experts to complete the land cover census for the UK's National Parks. Therefore, we need a repeatable, accurate and above all cost effective way of measuring land cover change across this large dynamic and varied landscape.

The Peak District National Park together with Cranfield University have started to explore ways to address this problem through automatic classification of land-cover and land-cover change. The main problems to overcome are the variation in the appearance of the same land-cover types across different images. Using sample data collected so far, and the latest high resolution aerial photography you will refine, train and test automated classification methods that are suitable for deploying automated classification at a landscape scale.

You will be working on this challenge full time with support and supervision where required.

TIN + DSG Spring 2022

Your responsibilities will include:

- Solving challenging business problems using advanced machine learning methods such as Deep Learning and quantitative analytics.
- Cleaning, aggregating and interpreting data in preparation for analysis.
- Extract information from aerial photography and satellite imagery.
- Explore the latest technology approaches as applied to this problem and contribute new methodological approaches where possible.
- Use open-source software tools and Python for image classification.

Expected Outcomes

- Development of, and implementation of a state-of-the-art model for assessing landcover in the context of replacing human visual interpretation.
- Production of an academic paper which documents the methods developed, results, discussion and recommendations.

The internship will also be instrumental in the preparation of a Data Study Group challenge that will follow on and expand the intern's work with the organisation. The intern will have the opportunity to represent the organization during the Data Study Group.

Supervision and Mentorship

You will be working with and supervised by the Peak District National Park and Cranfield University. National Park staff are experts in managing the landscape and can offer context and wider knowledge about the project through meetings and fieldtrips when required. You will also be directly supervised by Lectures in Remote Sensing from Cranfield University's Centre for Environmental and Agricultural Informatics, accessing their expertise and experience in this field.

Ideal Intern

The ideal candidate should have:

- An interest in physics, earth observation or image analysis
- Practical experience of coding in Python.
- Knowledge of machine learning modelling techniques and neural networks
- You should be experienced in using specialized machine learning tools e.g. Tensorflow, pytorch, sci-kit learn, etc.
- Must demonstrate capacity for reading, understanding and implementing new techniques in the field of machine learning as they emerge.
- Strong verbal/written communication & data presentation skills are recommended.

Internship Logistics

Start Date:	September 2022
Duration:	6 Months
Location:	Fully remote working, with potential for meetings at the Peak District National Park and Cranfield University
IT:	You will be expected to use your own PC, software will be open source
Remuneration:	£30,000 p/a pro rata

TIN + DSG Spring 2022

Benefits:

We offer a range of staff benefits including flexi-time, a great pension with employer contribution of 19.57%, 25 days holiday plus bank holidays and free parking at most of our offices. Please see our Total Reward Brochure to see what we offer.

Flexible Working:

A range of flexible working options would be considered.

Equality, Diversity and Inclusion:

The Peak District National Park Authority is committed to equal opportunities. We aspire to have a diverse and inclusive workplace and strongly encourage applicants from a wide range of backgrounds to apply and join us.

If you would like to discuss this, please contact David Alexander | Senior Data and Research Analyst on; david.alexander@peakdistrict.gov.uk

Plymouth Marine Laboratory Oceanographic AI Researcher TINDSG-007

#artificialintelligence #machinelearning #datastudygroup

About the Organisation

The [Plymouth Marine Laboratory](#) is an International Centre of Excellence in Marine Science & Technology and a Collaborative Centre of NERC. The research at PML contributes to the issues of global change, sustainability and pollution delivering solutions for national and international marine and coastal programmes. PML has an outstanding reputation at a national and international level for its capabilities in marine biogeochemistry and satellite Earth Observation.

Role Description and Responsibilities

Project title: Global decadal shifts in coccolithophore bloom distribution

Background Information

The only planktonic functional type that can be unambiguously detected from space are coccolithophores. This is because they are coated in thin calcite (calcium carbonate, CaCO_3) plates or liths. When these are shed into the water, they reflect incident light like tiny mirrors making them uniquely visible not just at the sea surface but also from satellite sensors, giving the water a turquoise, milky white appearance. Using satellite Earth Observation (EO) data these calcifying 'blooms' of coccolithophores are known to occur over large areas of the global oceans.

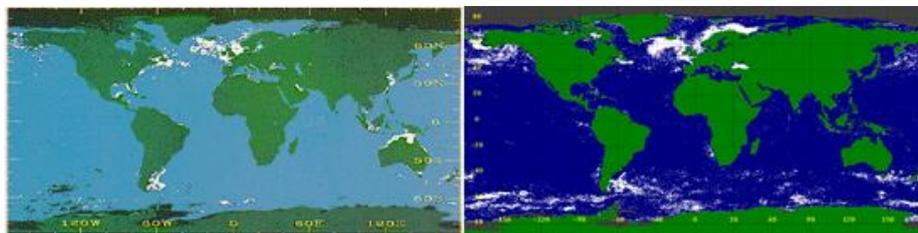


Fig 1. Presence (white) / absence of coccolithophore blooms from global analysis of satellite sensors CZCS (left) and SeaWiFS (right).

Primary Data Sources

We have previously constructed a global 40-year timeseries of these blooms (<https://doi.org/10.5194/essd-10-2043-2018>), but now need to describe their distribution and attribute any shifts in their patterns to other system drivers such as temperature, salinity, nutrients, circulation patterns, wind fields etc.

TIN + DSG Spring 2022

Analysis Methods Required

Solving the question of what the primary driving factors are behind global coccolithophore bloom distribution is a long-standing problem in marine ecology, which various investigators have sought to answer for the past 10 – 15 years. The main difficulties are that the datasets showing their distribution globally (consistently) are relatively recently available (last 5 years). We know that they are likely driven by changes in physical and chemical forcing, but finding these patterns and linkages is difficult. Therefore, this internship will investigate the changing patterns using skills in handling Big Data and employing statistical techniques within a framework of Artificial Intelligence and Machine Learning.

Team and Responsibilities

The intern will be part of the Marine Biogeochemistry and Observations group at PML and will work closely with the satellite EO group via the NERC Earth Observation Data Acquisition and Analysis Service (NEODAAS). They will be responsible for:

- Coding and implementing the various algorithms on the PML compute infrastructure (Python, Jupyter Notebooks)
- Keeping a high-quality record of their research and progress whilst at PML

Expected Outcomes

The outcomes of the internship will be a well-documented investigation into the shifts in coccolithophore bloom distribution in the global ocean. The intern will be expected to deliver high-quality, well commented and documented code showing the steps taken in the project including the investigations which were not ultimately successful.

The ultimate goal of this internship is a publication in a peer-reviewed journal, which the intern will be an integral part in delivering. If successful, this internship could help settle a 15-year debate on the changes to ocean ecology driven by climate change.

The internship will also be instrumental in the preparation for a Data Study Group challenge that will follow on and expand the intern's work with the organisation. The intern will have the opportunity to represent the organization during the Data Study Group.

Supervision and Mentorship

The intern will be directly supervised by Dr. Tim Smyth who is Head of Science for Marine Biogeochemistry and Observations at the Plymouth Marine Laboratory. There will also be collaboration and mentorship with other groups within the laboratory, in particular with NEODAAS.

Ideal Intern

The ideal candidate will have:

- Experience in handling multiple large datasets and implementing algorithms in Python/ Jupyter notebooks
- Familiarity with self-descriptive data formats such as netCDF, HDF and have skills with their manipulation within a coding framework.

TIN + DSG Spring 2022

- Experience of various AI / ML techniques, ideally using GPU technology; AI / ML techniques to include pattern recognition, neural networks and Bayesian reasoning
- Knowledge of Big Data standards and interoperability and have an ability to design a coding framework in order to address scientific questions
- Good organisational and time management skills and be well motivated to work independently as well as part of a wider team.
- An aptitude of turning their skills towards answering a long-standing question in oceanography and biogeochemistry.

Internship Logistics

The location for this internship will be at the Plymouth Marine Laboratory with the option of remote or hybrid working.

The start date will be early November 2022 with a duration of 4 months and will be full-time.

The salary is £30000 p/a pro rata.

The point of contact for technical queries: Dr. Tim Smyth (tjism@pml.ac.uk).

Sustrans Evaluation Officer TINDSG-008

#socialdatascience #datastudygroup

About the Organisation

[Sustrans](#) is the charity making it easier for people to walk and cycle. We connect people and places, create liveable neighbourhoods, transform the school run and deliver a happier, healthier commute. Sustrans works in partnership, bringing people together to find the right solutions. We make the case for walking and cycling by using robust evidence and showing what can be done. We are grounded in communities and believe that grassroots support combined with political leadership drives real change, fast.

From climate change and air pollution, to physical and mental health crises, the pressures on communities across our four nations are growing. There is no silver bullet but the work of Sustrans and our partners provides an essential contribution to tackling these challenges through encouraging and supporting modal change to emissions-free transport options. Sustrans' Research and Monitoring Unit (RMU) is the leading national centre of excellence for active travel monitoring, evaluation and research with 20 years of experience in providing independent and transparent evidence on walking and cycling.

Our Vision

A society where the way we travel creates healthier places and happier lives for everyone.

Our Mission

We make it easier for people to walk and cycle.

We have two strategic priorities



Our work leads to...

TIN + DSG Spring 2022



Role Description and Responsibilities

The intern will have a key role within the What Works team of Sustrans' Research & Monitoring Unit. The intern will report to the Head of Impact and working closely with our team of analysts.

RMU are specialists in designing and implementing approaches to assess what works in encouraging active travel behaviour and the effects of infrastructure interventions. The team designs and manages primary data collection but also have the added benefit of a Sustrans national evidence base of walking and cycling data. Our portfolio ranges from monitoring and evaluation of local projects to nationwide programmes with international significance.

We know that people, individually, benefit from walking and cycling, and society overall benefits when more of us are walking and cycling. But those benefits are not enjoyed equitably. Making it easier for people to walk and cycle leads to healthier places and happier lives, but not currently for everyone. Working closer to our values will mean that we must change our working practices to be truly for everyone rather than building where there is least resistance. We would like this change to be led by data and evidence as much as our ability to engage with different communities.

Our challenge is to bring data together that can help us understand the extent to which our work is equitable – for everyone. The questions we would like to investigate are:

- Whether the projects and infrastructure that we have delivered to date are geographically biased e.g. predominantly reaching affluent white communities
- How we can use data to inform future delivery to be more equitable

The data and information we have available for this include:

- GIS based location data of our projects and infrastructure
- Deprivation and Census data
- User surveys, including opinions from in-person surveys on cycle routes
- Automatic and manual count (e.g. traffic) data from our infrastructure projects

The intern's responsibilities may include:

Analysis:

1. Familiarisation with the datasets Sustrans has available, and the identification and collection of any relevant external datasets.
2. Developing methodology to allow inference on who is likely to have access to Sustrans projects and infrastructure, identifying biases, inequalities and usage patterns. This may include elements of spatial analysis, natural language processing, and time series analysis across heterogeneous datasets.
3. Making recommendations for how we can use data to inform future delivery to make

TIN + DSG Spring 2022

our projects more equitable, and considering “what would happen if” scenarios.

4. If required, work with colleagues and learning resources to get up to speed on the use of Arc GIS, R and Power BI (the main analysis tools Sustrans will use for this)
5. Prepare a Data Study Group challenge focused on further developing and expanding the methodologies and analyses scoped during the internship, and represent the organization during the event.

Sharing learning and communication:

6. Work closely with other analysts in the team, and communicate regularly about the project to the Research and Monitoring Unit
7. Ensure that any analysis and dashboard reporting is usable and accessible to colleagues
8. Talk to colleagues across the organisation including, but not exclusively, our colleagues with an Equality, Inclusion and Diversity remit and those working on the National Cycle Network (from engineering to volunteers)
9. Ensure that all material produced complies with the charity’s guidance on branding, tone of voice and key messages, positively contributing towards raising Sustrans’ profile

Expected Outcomes

For current and historic projects:

- key demographics of who lives close to projects and infrastructure
- a summary of key demographic reach by UK nations, regions and local authorities
- changes in key demographic reach over time
- if/where possible a comparison of survey data to Census data to see who we observe as using our projects/infrastructure vs who lives close

For future projects:

- the ability to update the analyses as new projects and infrastructure come online
- approaches for evaluating ‘what would happen if’ scenarios when formulating our projects

Project documentation including for example:

- A dashboard in Arc GIS or Power BI to display the results.
- Technical notes of processes and analysis developed
- Narrative of the task and outcomes to help bring the project to life to colleagues

Supervision and Mentorship

The intern will be part of the What Works team within Sustrans RMU and report to the Head of Impact.

Sustrans is committed to continuous improvement through learning and development. Each individual sets out learning and development through Sustrans’ personal impact and development process. The personal impact and development plans are reviewed regularly, and between meetings there is continuous dialogue between colleagues and their managers with monthly 1-2-1’s.

TIN + DSG Spring 2022

Ideal Intern

Sustrans aims to be a truly inclusive employer and welcome applications from people from all parts of the community, in particular from under-represented groups. Sustrans is committed to reducing inequality, valuing diversity and enabling inclusion.

	Application	Interview
Specific experience required		
Working with analysis tools such as R, Power BI and ArcGIS Online, or willingness to learn	✓	✓
Project Management		✓
Skills and Abilities		
Managing and working with large datasets, spatial and non-spatial	✓	✓
Analysing complex datasets and presenting findings	✓	✓
Writing and maintaining metadata for managing datasets		✓
Ability to work flexibly, as part of a team		✓
Excellent communications skills (general, written, verbal)	✓	✓
Simplifying technical documentation for a wide audience	✓	✓
Specific knowledge required		
Understanding of best practice in data analysis and dashboard reporting	✓	✓
An interest in and commitment to Sustrans' vision, mission and values	✓	✓

Internship Logistics

The intern will be able to utilise flexible hybrid working opportunities with the majority of Research and Monitoring Unit colleagues based in Edinburgh, Newcastle, Belfast, London and Bristol. There are office 'hubs' across the UK with colleagues based in Cardiff, Peterborough, Leeds, Nottingham, and Manchester. Our hybrid working policy welcomes the Intern to work in the way that suits them best, this could mean attending one of our main offices on most working days or only visiting an office occasionally.

TIN + DSG
Spring 2022

The start date will be in September 2022 with a duration of 6 months.

We would welcome interns working full-time or part-time but think candidates working a minimum of 3 days would suite the project best.

The salary is £30000 p/a pro rata.

The Rivers Trust

Machine learning approaches for Environmental data analysis

TINDSG-009

#machinelearning #datastudygroup

About the Organisation

[The Rivers Trust](#) is the umbrella body for the Rivers Trust movement, comprising 60 local charities delivering environmental improvements to rivers and their catchment areas throughout the UK and Ireland. The Rivers Trust has nearly 50 employees and has expanded rapidly in recent years. Find out more about the impact of our work in our [annual review](#).

With our expertise, reach and our member trusts' local knowledge, we work with farmers, government, businesses and local communities to deliver lasting environmental improvements. We bring together the people, knowledge, data and intelligence to help rivers thrive again; for us and future generations.

We strive to be an inclusive and collaborative movement and believe that by working together, we can make our shared vision a reality: wild, healthy, natural rivers, valued by all.

Role Description and Responsibilities

You will be part of an exciting transformational change project called the Catchment Systems Thinking Cooperative (CaSTCo) which is a three-year Ofwat-funded innovation partnership between The Rivers Trust, 12 water companies and 11 other academic, commercial and NGO partners.

The CaSTCo project is launching this summer and will revolutionise the way crucial data about England and Wales' water environment is gathered and shared, in particular on the health of the nation's rivers. Only 14% of rivers in England are in 'good ecological health' – one of the worst records in Europe, yet we lack the density of information needed to fully understand the nature and sources of intermittent problems such as pollution, or the effectiveness of Nature Based Solutions at delivering multiple beneficial outcomes such as flood protection and water quality improvements. We need to fill evidence gaps with enough certainty to drive action and, crucially, to bring together the partners and funding streams which could tackle these complex problems. This project will start to build a more robust evidence base for tackling environmental challenges by developing a national framework of standardised monitoring tools, training and data platforms. A major focus of the project is that it will directly engage and empower local communities in gathering local evidence in eight demonstrator catchments, feeding this information into decision support tools and decision-making processes.

The CaSTCo project will upskill and mobilise citizen scientists to collect environmental data alongside wider deployment of 'continuous' monitoring sensor networks. Your role will be to streamline the interpretation and extraction of insight from this sensor data so that it can be used and applied by all project stakeholders. Through the project partners you will have

TIN + DSG Spring 2022

access to a team of water quality monitoring experts that can provide interpretation of the data, help define the scope and test the machine learning approaches which you develop.

You will have the opportunity to work with businesses, environmental NGOs, government, local communities and campaign groups to support greater use and understanding of data for detecting pollution sources and events, and for underpinning sustainable catchment management decision-making.

The Rivers Trust has secured access to a number of data sources for use on this project, including:

- Continuous water quality sensor data from the Environment Agency's National Water Quality Instrumentation Service – data variables include 8 key water quality parameters (temperature, turbidity, conductivity, pH, dissolved oxygen, chlorophyll, blue-green algae and ammonium)
- Water chemistry data from the Environment Agency's water quality archive – over 58 million rows of data from the last 20 years of spot sampling undertaken by the EA
- Continuous water quality data from a network of simple sensors (measuring temperature, water level and conductivity) deployed at high spatial resolution across a river catchment (7 sensor nodes in a catchment of 14 sq km area). This network has specifically been deployed to collect a 'training' data set to enable greater understanding of dynamic water quality characteristics at the catchment scale
- Continuous water quality data from water quality sondes and tryptophan-like fluorescence sensors (an indicator of bacterial contamination)
- Rainfall and river level data from the Environment Agency's flood monitoring network
- Citizen Science data comprising 8000 surveys across the Westcountry between 2016 and 2022. The data set is still growing and includes records of wildlife sightings, pollution sources, problem plants, water quality and flow conditions – with accompanying photographs

You will work within The Rivers Trust Technical Team – a team of around 10 technical experts in the fields of spatial data analysis and visualisation, environmental modelling and monitoring. We do not currently have machine learning expertise within the team, so you will help build our understanding of the potential for application of machine learning and artificial intelligence approaches to environmental data science. Members of the Technical Team will be working on the CaSTCo project – liaising with the catchment demonstrators, developing data collection and monitoring plans and protocols for sharing data from diverse sources. You will work closely with the team to integrate your work in to the CaSTCo programme.

Expected Outcomes

This internship offers a unique opportunity to help shape the future of water quality and environmental data analysis, interpretation and collection in England and Wales. Your work will be aligned with an innovation project which is being co-designed by the project partners, so you will have the opportunity to shape and design the processes and specific deliverables in a highly collaborative environment. You will be supported by leading national water quality and environmental data experts in interpreting and deriving insight from a range of archived and live data sets and your work could influence the direction of future environmental monitoring strategies.

Some of the outcomes which we anticipate from this internship, the Data Study Group and the CaSTCo project include:

- Improved understanding of the potential and limitations of machine learning and artificial intelligence approaches for water quality / environmental data analysis and

TIN + DSG Spring 2022

interpretation.

- Improving the data literacy of communities and NGOs and enabling better understanding of the multiple factors affecting river health in space and time and the interdependencies between these factors.
- Better information for swimmers and recreational river users about dynamic changes in water quality and the factors that influence whether a river is fit to swim in.
- Application specific indicators derived from multiple data sources that help to identify pollution events and characterise catchment response to rainfall events i.e. automatic discrimination between point source pollution and rainfall-driven run off events
- Automatic flagging of suspect data by learning to recognise common features that are indicative of instrument failure or reduced performance (e.g. sensor biofouling, drift, etc.)
- Identifying long term trends against a very 'noisy' data background – continuous water quality data is subject to wide variations over a range of timescales – diurnal variations, weather and seasonally driven variations can obscure long term trends or changes due to land use, new developments or waste treatment asset performance
- Sensor based 'classification' of river health – typically the health of rivers has been summarised using spot testing results which, whilst subject to rigorous laboratory analysis is often collected at very poor temporal and spatial resolution. We would like to investigate the potential to provide an alternative approach to classifying some aspects of river health using high temporal resolution data from continuously recording water quality sensors

The impact of these outcomes will help to transform the use of sensors in the water environment, and we anticipate that this will increase demand and in turn influence sensor manufacturers to focus research and development into the areas of lower cost sensors and technologies that are deployable by local communities and citizen scientists.

The internship will be instrumental in the preparation of a Data Study Group challenge that will follow on and expand the intern's work with the organisation. The intern will have the opportunity to represent the organization during the Data Study Group.

Supervision and Mentorship

You will be jointly supervised by The Rivers Trust's Technical Director and Technical Lead for the CaSTCo Project. You will also be mentored by experts from partner organisations, including member Rivers Trusts, the Environment Agency, water companies and academic and commercial partners. We will also seek opportunities to engage you in some of the innovative machine learning research being developed by our partners.

Ideal Intern

We are looking for a skilled data or computer scientist with expertise in the design and application of machine learning approaches who is keen to work in a dynamic NGO environment. You will be passionate about using your skills to tackle some of the most pressing environmental challenges of our time and will relish the opportunity to engage with an interdisciplinary team to create innovative solutions to data science questions.

You will bring both technical skills and knowledge and an enthusiastic, people-oriented approach to this role. You will be willing and able to shape the approach to this challenge,

TIN + DSG Spring 2022

liaising with project partners to build your understanding of the project and the gaps in knowledge which it is aiming to tackle. You will use your research skills and machine learning knowledge to explore the datasets, extract features and design and test a range of machine learning approaches, including supervised and unsupervised learning, deep learning and uncertainty quantification. You will liaise with data scientists, sensor manufacturers, environmental decision makers and project stakeholders to design and scope an appropriate Data Study Group challenge which can contribute to the overall aims of the CaSTCo project.

You will have great communication skills and a willingness to share your knowledge and build data literacy among a wide range of stakeholders that you will encounter on the project. The Rivers Trust has a strong open data and knowledge sharing ethos, and we will support you to produce open research outputs from your work through conferences, data impact case studies or peer reviewed publications.

Key Qualifications:

- Strong data science, machine learning and coding skillset is essential
- Experience with time series analysis
- Excellent problem solving, critical thinking, and communication skills
- Driven, self-motivated and highly organised
- Comfortable with dynamic, innovative and creative challenges
- Environmental science, especially water quality knowledge would be desirable

Education: Pursuing a PhD in computer science, data science or related technical field.

Internship Logistics

Salary: £30,000 per annum pro rata

Length of Internship: Six months, fixed term. Full time (37.5 hrs per week with flexible working arrangements available). Part-time will be considered for exceptional candidates.

Start date: September 2022

Location: UK home based with some travel required for occasional face-to-face meetings in England and Wales.

For technical queries, please email: michelle.walker@theriverstrust.org