# MAchine learning for environmental Time Series (MATS)

Coordinator: Romain Tavenard
LETG / Université de Rennes

| | |
|---|---|
| Acronym | MATS |
| Titre du projet | Apprentissage Statistique pour les Séries Temporelles Environnementales |
| Proposal title | MAchine learning for environmental Time Series |
| Type of research | Basic Research |
| Grant requested | 215 000 € |
| Project duration | 48 months |
| Coordinator | Romain Tavenard - LETG / Université de Rennes |
| Type of project | JCJC (Young researcher) |
| Scientific committee | 23: "Données, Connaissances, Big data, Contenus multimédias, Intelligence Artificielle" |
| Keywords | Time series analysis ; machine learning ; remote sensing of environment |

## Contents

## ABSTRACT OF THE PROJECT

A huge trend in recent earth observation missions is to target high temporal and spatial resolutions (*e.g.* SENTINEL-2 mission by ESA). Data resulting from these missions can then be used for fine-grained studies in many applications. In this project we will focus on three key environmental issues: agricultural practices and their impact, forest preservation and air quality monitoring.

Based on identified key requirements for these application settings, MATS project will feature a complete rethinking of the literature in machine learning for time series, with a focus on large-scale methods that could operate even when little supervised information is available. In more details, MATS will introduce new paradigms in large-scale time series classification, spatio-temporal modelling and weakly supervised approaches for time series. Proposed methods will cover a wide range of machine learning problems including domain adaptation, clustering, metric learning and (semi-)supervised classification, for which dedicated methodology is lacking when time series data is at stake. Methods developed in the project will be made available to the scientific community as well as to practitioners through an open-source toolbox in order to help dissemination to a wide range of application areas. Moreover, the application settings considered in the project will be used to showcase benefits offered by methodologies developed in MATS in terms of time series analysis.

**Changes that have been made in the full proposal compared to the pre-proposal**

Scientific program of the MATS project has not significantly evolved since the pre-proposition (though tasks are described in more details in the current document), and we have added the organization of an international workshop on machine learning for time series in order to better disseminate the outputs of the project and foster scientific collaborations on MATS topics of interest. Organization of this workshop, which was not planned in the short proposal, induces additional costs for the project. Also, a consortium agreement, which was not planned in the pre-proposal, is considered in this version to settle intellectual property issues. However, considering these two modifications, the budget increase remains within the 15%-margin authorized by ANR.

**Summary table of persons involved in the project**

| Affiliation | Name | First name | Current position | Role & responsibilities in the project | PM |
|---|---|---|---|---|---|
| LETG | Tavenard | Romain | *Maître de Conférences* | Coordinator. Involved in all tasks. time series; machine learning; remote sensing | 24 p.m |
| | Corpetti | Thomas | *Directeur de Recherche* | Involved in tasks 2.3, 3.1, 4.3 time series ; remote sensing | 6 p.m |
| | Alvarez | Emilien | *Stagiaire*, PhD student soon[1] | Involved in task 4.1 remote sensing | 6 p.m |
| | Oszwald | Johan | *Maître de Conférences* | Involved in task 4.2 remote sensing | 4 p.m |
| IRISA Linkmedia | Malinowski | Simon | *Maître de Conférences* | Involved in tasks 1.1, 3.2 machine learning ; time series | 12 p.m |
| IRISA Obelix | Chapel | Laetitia | *Maître de Conférences* | Involved in tasks 1.2, 2.1, 2.2, 3.1, 3.3, 4.2 machine learning | 12 p.m |
| | Friguet | Chloé | *Maître de Conférences* | Involved in tasks 2.1, 2.2, 3.1 machine learning | 9 p.m |
| | Gloaguen | Pierre | Post-doc | Involved in tasks 2.1, 2.2, 3.1, 4.2 machine learning | 6 p.m |
| | Courty | Nicolas | *Maître de Conférences* | Involved in task 3.3 machine learning ; optimal transport | 4 p.m |
| | Vayer | Titouan | PhD student | Involved in task 3.3 machine learning ; optimal transport | 6 p.m |
| University of East Anglia | Bagnall | Anthony | Senior Lecturer | Involved in tasks 1.2, 3.1, 3.2, 4.1 time series | 12 p.m |
| | Lines | Jason | Lecturer | Involved in tasks 1.2, 3.1, 3.2, 4.1 time series | 12 p.m |
| Monash University | Petitjean | François | Senior Researcher | Involved in tasks 1.2, 3.2, 4.1 machine learning ; remote sensing | 6 p.m |

[1] Emilien has obtained a PhD scholarship from Région Bretagne and Université de Rennes 2. His PhD work will take place at LETG-Rennes.

# 1 Proposal's context, positioning and objective(s)

## 1.1 Objectives and scientific hypotheses

A huge trend in recent earth observation missions is to target high temporal and spatial resolutions. As an example, SENTINEL-2 mission by ESA delivers data for the whole planet at a 10m to 60m spatial resolution (depending on the spectral bands considered) with a revisit time of 5 days. Data resulting from such missions hence offers very promising perspectives for a wide range of applications. In order to turn this data into knowledge, it becomes crucial to develop methodologies that are able to deal with **large volumes** of **temporal data**. In this project, we will focus on the monitoring of environmental systems using satellite imagery. In this context, we have identified three key application settings that will be tackled by the project.

First, a focus will be given towards the **monitoring of agricultural practices**. To do so, we will develop in this project efficient tools for the identification of such practices, which is the necessary basis for later analysis of its impact on biodiversity and the environment in general [1]. Several time-series-specific machine learning tools will be considered. Indeed, for each location on the land cover map to be generated, a time series can be extracted from series of satellite images, making the land cover (or land use) recognition task a time series classification task. Moreover, the diversity of acquired data (from year to year, from modality to modality, or from temporal resolution to temporal resolution), together with the high cost of data labelling indicates a high interest for domain adaptation techniques that could benefit from the availability of multiple data sources. More generally, semi-supervised approaches that would leverage on the vast amount of available unlabelled data and make the most out of the few labelled samples would be of great help in this context.

Second, we will provide tools to better understand the phenomena at stake in **forest degradation**. It is well documented that forests are "critical part of the solution to the unprecedented demand for water, food and energy" (source: United Nations' website[2]). In this context, we will provide methods to quickly identify the presence of various degradation / regeneration patterns from satellite image time series. Also, anomaly detection tools will be used to identify potential forest dieback cases in large datasets. Due to the high human cost required to get reliable labelled data coupled with the high number of tree species at stake, which results in a high variability in vegetation index profiles, unsupervised methods will be preferred for this application case.

Last but not least, we will target the **pollution monitoring** issue. Air quality is at the moment one of the main issues for urban planning. Prospective scenarios expect indeed that in 2050, two third of the world population will live in cities, as stated by a UN report in 2014.[3] Ensuring a reliable quality of life in urban environments is then of prime importance and air pollution monitoring and forecasting focuses number of researches at the moment. In this project, efficient tools for understanding urban flow patterns (directly connected to air quality) as well as new and fast simulation tools will be designed.

These application settings shed light on the dramatic need for tools to perform time series classification at large scale and propose accurate time series models so as to get a better understanding of the underlying phenomena at stake. Such tools should be easily accessible to a wide community of practitioners. Developed tools should be able to (i) deal with the large volume of data at stake, (ii) adapt to scenarios in which the labelling cost (in terms of human resources) is especially high, and (iii) take spatial coherence into account to obtain better insights into the data.

Regarding these needs, the MATS project features substantial breakthrough in the following areas:

1. large-scale time series classification (*cf.* task 1 description below);
2. spatio-temporal modelling of geo-localized time series (*cf.* task 2);
3. weakly supervised approaches for time series (*cf.* task 3).

Also, the need for efficient tools that could be used by practitioners will be covered by the development of the `tslearn` toolbox, and application of the developed methodology to real settings described above will be driven in collaboration with experts from the remote sensing community.

[2] https://news.un.org/en/story/2018/03/1005561
[3] https://esa.un.org/unpd/wup/

## 1.2 Originality and relevance in relation to the state of the art

**Positioning versus the state-of-the-art**

Before providing details on the key ambitious and innovative objectives of MATS, we provide in what follows a state-of-the-art on machine learning methods for time series.

In the time series classification setting, which is by far the most widely considered in the literature, two families of methods can be distinguished: those that focus on the design of dedicated similarity measures and those that rely on the extraction of features to embed time series in metric spaces in which standard machine learning tools can be considered. The most well-known dissimilarity measure for time series is the dynamic time warping (DTW) [2]. DTW finds a minimal cost alignment path between pairs of time series, which makes it invariant to temporal distortions. Other alignment-based metrics have been proposed to extend DTW into a positive definite kernel [3] or a differentiable similarity function [4], making it easier to embed at the core of standard machine learning techniques such as kernel methods or neural networks. The feature-based family of methods includes works relying on hand-crafted features [5, 6] as well as learning-based approaches, among which the shapelet model plays an important role. Shapelets are time series sub-sequences with a high discriminative power [7]. In this framework, classification is done with respect to the presence or absence of given shapelets in the time series. Shapelets can either be searched for in the time series [7], drawn at random [8] or learned [9]. The latter case is similar in spirit to neural network approaches, for which both convolutional [10–12] and recurrent [13] variants have been investigated for time series classification tasks. Important enough to be noticed, most of these works make use of rather shallow structures (*e.g.* the *Learning Shapelets* [9] model only uses one hidden layer), when compared to models used in the computer vision community for example.

Much fewer attention has been paid up to now to other machine learning tasks such as clustering, domain adaptation or metric learning. In the unsupervised context, some works are targeted towards the use of a specific metric at the core of standard clustering algorithms [4, 14, 15], while others aim at learning adequate embeddings [16]. The *Learning DTW-Preserving Shapelets* (LDPS) model [17] proposed by members of the MATS project somehow acts as a synthesis between both options, in which an embedding is learned so as to mimic a target metric. The latter can be seen as a metric learning approach in which the target metric is explicit, while it is implicit in [18]. The semi-supervised setting has retained little attention up to now and proposed methods have focused on self-labeling techniques [19, 20]. Concerning domain adaptation, a first effort has been done by members of the MATS project so as to release datasets for the evaluation of domain adaptation strategies for remote sensing time series [21], but no time-series specific method has been proposed up to now, to the best of our knowledge. Time series modelling has received more attention. Works developed in this context include fine-grained yet monomodal continuous time models [22], variants of auto-regressive models [23, 24] that assume evenly-spaced observations, or hierarchical topic models [25] that rely on quantized temporal information.

MATS project will consist in a complete rethinking of the approaches for machine learning with time series through the proposition of **large-scale time series classifiers**, the development of new paradigms in **weak learning or unsupervised settings for time series** and the proposition of ground-breaking generic **spatio-temporal models**.

In more details, based on previous works of members of the proposal [12, 17], supervised approaches that bridge the gap between shapelet models and neural networks will be explored with the intention to allow for deeper structures to arise in the time series community so as to address truly-large-scale datasets. The development of models that require little to no supervision will also be an important part of the project, in order to tackle domain adaptation or metric learning settings, for example. To do so, we will rely on the expertise of members of the project on time series clustering [14, 17], domain adaptation [21] and temporal topic models [26]. From an application point of view, proposed methods will be used to help solve environmental issues using modern remote sensing data sources such as the SENTINEL-2 earth observation mission. This application context is in line with the methodological developments proposed in the project since it corresponds to cases where large amounts of spatio-temporal data are generated while supervised information is lacking.

**Positioning versus national and international initiatives**

Many research activities dealing with time series are carried out in France. In particular, a couple of ongoing ANR projects focus on the development of dedicated methodology for time series data. The TIMES project ("High-performance processing techniques for mapping and monitoring environmental changes from massive, heterogeneous and high frequency data times series")[4] shares use cases with MATS, but it is oriented towards the specific case of change detection, while MATS covers a wider scope of machine learning paradigms. The LOCUST ("Representation Learning for Modeling Rich Dynamic Interaction Traces") project[5] is focused on representation learning for the analysis of navigation traces from the Web, which biases proposed methods towards a specific sub-class of time series (*i.e.* paths in the Web graph) that does not match MATS requirements. The PIA-funded IKATS project[6] deals with the analysis and visualization of time series data with a focus on prediction techniques for industrial applications. Finally, the ERC-funded SEDAL ("Statistical Learning for Earth Observation Data Analysis") project[7] is similar in spirit to MATS, as it relies on the development of novel machine learning tools with application to remote sensing data, yet it does not focus on the development of time-series-specific tools.

### 1.3 Methodology and risk management

The MATS project is ambitious both in terms of methods and applications. In order to ensure that these applications can be tackled, we will take care that necessary methodology is successfully introduced. To do so, we propose, in each methodological task described below, low risk alternatives to the high risk-high gain preferential research tracks. Section 2.3 details levels of risk and associated alternatives for each task.

Dissemination of the project outputs will be done through publications in machine learning and remote sensing areas. A python toolbox created by the project coordinator will also be used to make the developed methods easily accessible to the research community and its download statistics will serve as an additional indicator of success for the project. Finally, the organization of an international workshop on the topics of the project will also be an important point of our dissemination strategy, as well as a unique occasion of building new collaborations with international researchers of the field.

## 2 Project organization and means implemented

### 2.1 Scientific coordinator and its consortium / its team

MATS project will be led by Romain Tavenard. Romain gratuated from École Centrale de Lyon and École Normale Supérieure de Cachan. He has spent one year at the Centrum voor Wiskunde en Informatica (CWI, Amsterdam, The Netherlands) and two years at the Idiap Research Institute (Martigny, Switzerland). Romain obtained his PhD in 2011 and he is currently an assistant professor at the Université de Rennes since September 2013. He has been granted the *Prime d'Encadrement Doctoral et de Recherche* in 2017. His research activities take part in the LETG laboratory, where his main research interests are machine learning for time series data, with application to environmental time series. Romain is the creator and main developer of the `tslearn` python toolbox that will be widely used and extended in the context of this project. Romain features a strong experience in collaborative research projects, with past and present responsibilities as work package leader or partner coordinator. He will be involved in all tasks of the project.

The project will lead to collaborations with researchers with complementary backgrounds from time series analysis (Anthony Bagnall, Jason Lines, University of East Anglia – UK, Simon Malinowski, IRISA-Linkmedia), machine learning/statistics (Laetitia Chapel, Nicolas Courty, Chloé Friguet, Pierre Gloaguen, Titouan Vayer, IRISA-Obelix) and remote sensing (Emilien Alvarez, Johan Oszwald, Thomas Corpetti, LETG)

[4] http://www.agence-nationale-recherche.fr/Project-ANR-17-CE23-0015
[5] http://www.agence-nationale-recherche.fr/Project-ANR-15-CE23-0027
[6] http://www.lemagit.fr/article/IKATS-le-projet-francais-de-Big-Data-industriel
[7] https://sedalproject.wordpress.com/. Project description states that: "Despite the many successful results and developments, there are still strong limitations for the general adoption of machine learning algorithms for predicting and understanding EO data."


Figure 1: Gantt chart of the project.

fields. Furthermore, the joint expertise of François Petitjean (Monash University – Australia) in machine learning and remote sensing will be a definite asset for a project such as MATS. Some of these collaborations will be the continuation of existing ones (*e.g.* with IRISA members) while others will be initiated in the context of the MATS project (*e.g.* with researchers from University of East Anglia and Monash University).

Additionally, a PhD student with background in machine learning will be hired to work on tasks 1 and 3 and an engineer will be recruited to ensure integration of the developed methods in the `tslearn` toolbox.

## 2.2 Means of achieving the objectives

MATS consists in one management task and 4 scientific tasks that are detailed in the remaining. Among the scientific tasks, tasks 1 to 3 will be devoted to novel methodological developments, while task 4 is focused on the application of proposed methods to help solve important environmental problems. Figure 1 shows the Gantt chart of the project. The application task will have two phases: an early phase will be focused on the analysis of the data at stake and the application of the developed methodologies to this data will start after 18 months. Besides this data-centered activity, most of the energy at the beginning of the project will be devoted to implement the fundamental tools that will be used throughout the project. This Gantt chart illustrates to some point the degree of independence of the different methodological subtasks which will be considered in parallel by the project members.

### 2.2.1 Organization by task

**Task 0: Project management, dissemination and communication**

*People involved:* all

*Task 0.1: Communication, dissemination and related tools*

The objectives of this task are the early implementation of the project website, intranet and adapted collaboration and communication tools. Then it will be devoted to publication issues, and dissemination of the expected results. Dissemination strategy for the MATS project will rely on two actions. First, the development of the `tslearn` toolbox will favor a wide adoption of methods produced in the context of the project. Indeed, `tslearn` has already attracted attention from researchers in the machine learning domain as well as from applicative research domains, and we plan to rely on this constantly growing community to get publicity for MATS methods. Second, dissemination of MATS will also take the form of the organization of an international workshop on machine learning for time series. This workshop will either be part of a major conference or an independent workshop, and should be organized during the second half of the project in order to capitalize on MATS developments.

*Task 0.2: Project management*

The overall objective of this task is to create and maintain the conditions needed to achieve the objectives and deliverables of the project as described. Specific objectives are to ensure that scientific, technical and managerial decisions are committed for the success of the project, make sure that reporting and budget reviews are as specified in the proposal, coordinate knowledge management activities and ensure that the project outcomes are disseminated to researchers.

Two plenary meetings will be organized every year, to ensure the coordination of the project and review the general advances. Phone and web conferences or technical meetings will also be organized to address daily concerns. Issues regarding the intellectual property of the different outputs of the project will be detailed in the consortium agreement, whose editing will be coordinated by SATT Ouest-Valorisation (Technology transfer acceleration company) and in collaboration with the juridical offices of the other institutes participating to this project. In particular, this agreement will cover: (i) the intellectual property of previous knowledge and (ii) the intellectual property of the developments issued from the project.

**Task 1: Time series classification**

This task aims at both defining new local features for time series and better understanding relationships between existing representations and convolutional neural networks.

*Task 1.1: Feature localization*

*People involved:* Simon Malinowski, Romain Tavenard

A wide range of state-of-the-art feature-based time series classification models ignore feature localization information [6, 27]. We have shown in [28] that taking such information into account in a match-kernel framework significantly improves performance on a wide range of time series classification problems.

In this task, we will investigate extensions of this work to the shapelet framework [27]. In this framework, shapelets are learnt [9] (or searched for [7], or randomly drawn [8]) such that similarities between a time series and shapelets from a set can be used as a feature vector representing the time series.

A first option we will consider consists in incorporating localization information in this feature vector. To do so, we will use the random shapelet setup [8], in which shapelets are drawn from the training time series. In this setup, it is hence possible to retain, for a given time series, not only its distance to each of the drawn shapelets, but also the matching positions. Based on that representation, a feature selection scheme should be defined that would take into account the pairwise nature of the variables (for each shapelet, two variables are derived: one for the distance, the other for the localization). We will derive an adaptation of the sparse-group lasso framework [29] in which there would exist an ordering for variable selection inside a group. Indeed, it would make no sense to keep localization information into account while removing the associated distance, but the contrary should be possible (which would let the model learn whether the localization information is discriminant or not, on a per-shapelet basis). Such a structured variable selection framework could be beneficial to other uses cases outside the time series classification community.

Another related approach based on random shapelets will make use of the Distance to Kernel and Embedding (D2KE [30]) framework. The idea here will be to draw random shapelets and derive a kernel from any similarity measure between time series and shapelets. For example, elastic similarity measures could be considered to assess for the quality of a shapelet match [31], instead of standard euclidean distance. The obtained kernel could then be refined by taking temporal information into account using a convolutional kernel as in [28].

*Task 1.2: Bridging the gap between shapelet models and convolutional neural networks*

*People involved:* Anthony Bagnall, Laetitia Chapel, Jason Lines, François Petitjean, Romain Tavenard

More and more studies (including works by members of the MATS team [12, 17]) point out similarities between shapelet models and convolutional neural networks (CNNs). However, no clear conclusion about setups in which one model would be likely to outperform the other has been proposed, to the best of our knowledge. We believe a more thorough comparative study could be key to better understand the pros and cons of each approach and how the vast literature about neural networks could feed research about shapelet-based models.

Such a study will be the basis for this task and shall lead to a better understanding of why, in the time series community, shallow models tend to be preferred to deeper ones as those widely used in the computer vision community for example. One possible reason for that fact is that there is a lack, in the time series community, of large-scale datasets (large number of instances, long time series, large number of classes) that could highlight the strength of deeper models. In that context, we will push an effort towards the design of better regularization strategies for shapelet models, in order to be able to train deep shapelet models even on moderate-size datasets. Also, attention will be paid to the development of feature visualization techniques to better understand the multi-scale analysis capabilities offered by such deep shapelet architectures. Moreover, the lack of large datasets will be tackled in the frame of this project (*cf.* the “Data means” paragraph in Section 2.2.2 below).

Vulnerability of shapelet models to adversarial strategies will also be studied in this task, as well as the development of dedicated counter-measures. The latter would include standard strategies such as adversarial training [32] or distillation [33]. However, more time-series specific strategies will also be considered. For example, let us consider the early classification framework, which has recently attracted some research attention, including from members of the MATS project [34]. In such a setting, an attacker could get reward for delaying the time of correct classification, which should be taken into account when designing adequate counter-measures.

The models developed in this task will result in powerful vectorial representations for time series, which could then be used as an input for standard methods for a wide variety of machine learning tasks, with an emphasis on scalable methods that could deal with large volumes of data (*cf.* task 4 requirements below).

**Task 2: Time series modelling**

This second methodological task is focused on the design of generative models for time series. In this context, two families of models will be tackled: topic models and generative adversarial networks. Finally, an extension to spatio-temporal models will be discussed.

*Task 2.1: Continuous time topic models*

*People involved:* Laetitia Chapel, Chloé Friguet, Pierre Gloaguen, Romain Tavenard

In this task, we will investigate the use of topic models (on which members of the project have already worked [26]) for the analysis of temporal data. Application of such models to continuous-time and/or irregularly sampled data is usually done through quantization in both feature space and temporal dimension [25]. Such an approach, though easy to implement, suffers from a major drawback, which is that words are learned blindly with respect to the temporal motifs. In practice, words are formed via clustering by optimizing on a quantity (*eg.* inertia when $k$-means is used) that is unrelated to the quality of the final topic model to be learned. Moreover, the learned model is fully dependent on the chosen quantization, and difficult to compare with results obtained from a different quantization.

To tackle this limitation, we will consider two options. First, an *a posteriori* refinement step could be performed after the sequence of observations has been segmented into motif occurrences by a time-sensitive topic model such as [25]. This refinement step could be a way to incorporate expert knowledge through, for example, stochastic differential equations, a class of continuous-time models that has gained popularity in ecology (for instance, for the modelling of GPS data [22]). Second, we will develop a new family of topic models based on Dirichlet processes that integrate quantization in the model instead of seeing it as an independant pre-processing step.

Such models could be used to define normality models in applications related to anomaly detection. In this context, developed models should allow for online anomaly detection to better cover application needs as specified in task 4. Also, in order to allow for the use of such models in large-scale applications, variational inference will be considered as the preferred learning paradigm, which would constitute an important improvement over the state-of-the-art that relies on Gibbs sampling to learn such motifs [25].

*Task 2.2: Temporal generative adversarial networks*

*People involved:* Laetitia Chapel, Chloé Friguet, Pierre Gloaguen, Romain Tavenard

Another important family of generative models is that of the generative adversarial networks (GAN) [35]. These models rely on a two-player game in which a generator network and a discriminator one are involved with concurrent learning objectives. Though GANs are a particularly hot topic at the moment, little attention has been paid to temporal variants (except [36] for the case of videos, which does not fully match requirements for our application settings).

In this context, we will consider two options. First, a convolutional variant will be developed that will build on powerful models from the computer vision community [37]. In this context, insights obtained from task 1.2 on the connections between CNNs and time-series specific models shall be of great help. Second, a more ambitious formulation would be to let the generator learn to generate sequences of arbitrary length (which fully-convolutional models will not do). To do so, we will get inspiration from SEQ2SEQ models [38] and propose an extension of these models to real-valued data.

In all cases, as discussed in task 4.1, such models could be used to generate realistic evolution profiles for land cover applications. In this context, we will take care to design models that do not drastically suffer from mode collapse, for example by taking inspiration from [39].

Finally, specific attention will be paid to the design of models that could to deal with irregularly sampled data. Indeed, in many of the applications targeted in this project, data is acquired at irregular time intervals (due to atmospheric conditions perturbing the acquisition of satellite images, for example) and this setup, which is rather common in the time series domain, is not currently tackled in the neural network community, to the best of our knowledge.

*Task 2.3: Taking spatial arrangement into account*

*People involved:* Thomas Corpetti, Romain Tavenard

Models that MATS members will develop in the context of tasks 1 and 2 are oriented towards dealing with the temporal nature of the data at stake. However, remote sensing applications such as those targeted in this project often (if not always) rely on geo-localized time series. We plan to leverage on this spatial information in two ways.

First, spatial redundancy can be used as a denoiser. In that case, a spatial pooling layer could be appended at the output of models developed in tasks 1 and 2 so as to spatially robustify predictions at the object-level (*eg.* at the parcel level for the agricultural monitoring application considered in task 4.1).

Second, many of the applications targeted in this project have underlying spatio-temporal behaviours that one should model in order to better understand the observed phenomenons. For example, for flow applications as air pollution monitoring (*cf.* task 4.3), various classes of spatio-temporal patterns (related to the strength and dilation of rotating structures) are directly connected to the dispersion of pollutants and hence the local air quality. Currently, the modelling of these dynamical patterns is usually performed through proper orthogonal decompositions (POD) which are temporal extensions of principal component analysis [40]. In this project, a spatial extension of temporal generative adversarial networks developed in task 2.2 will be proposed to design spatio-temporal basis functions of local patterns of urban flows.

**Task 3: Weakly supervised learning**

Tasks 1 and 2 above are focused on fully-supervised (task 1) or fully-unsupervised (task 2) settings. We now turn our focus on mixed scenarios in which supervised information is available only for a subpart of the data or does not come in the form of traditional class labels.

*Task 3.1: Semi-supervised learning*

*People involved:* Anthony Bagnall, Thomas Corpetti, Laetitia Chapel, Chloé Friguet, Pierre Gloaguen, Jason Lines, Romain Tavenard

One important aspect of this project is to derive methods that can deal with as little supervised information as possible. This is motivated by (i) the fact that there is a clear lack of such methods in the state-of-the-art of time series analysis and (ii) application needs are widely supporting the development of such models (*cf.* task 4 for more details).

In this task, we will focus on the joint use of large amounts of un-annotated data and few labelled data. To do so, we will use temporal GAN models developed in task 2.2. More specifically, such compound models include a discriminator that is trained to distinguish between artificially generated data and real data. Such a model can be trained at large scale, as it does not require any labelled data. Hence, it should be able to learn very powerful features to efficiently represent time series data at stake. Based on this model, we will use fine-tuning to re-train its classification layers for the classification task at hand. This fine-tuning step will take full benefit of the little annotated data available.

Another approach that will be considered for this task is to use unsupervised models such as the *Learning DTW-Preserving Shapelet* (LDPS) one [17] and alter its cost function so as to take into account both metric reconstruction constraints and class information for labelled time series. This will result in an optimization scheme that alternates between unlabelled time series pairs for which the objective is to approximate a target metric (*eg.* the dynamic time warping one) and labelled time series for which a standard classification loss can be used.

*Task 3.2: Metric learning*

*People involved:* Anthony Bagnall, Jason Lines, Simon Malinowski, François Petitjean, Romain Tavenard

As stated above, *Learning DTW-Preserving Shapelet* (LDPS) has shown the feasibility of unsupervised learning of shapelet models [17]. The method, developed by members of the MATS project, consists in learning a representation that could mimic DTW and proved very successful when used to embed time series in a metric space in which clustering (euclidean $k$-means in our case) could be performed.

A reverse approach could be used when little amount of supervision is found in the form of must-link and cannot-link constraints [41, 42]. In such a case, one would not have a target metric to mimic but she would rather learn a metric that would be coherent with the expressed constraints.

In that context, we will consider the use of variants of the siamese model presented in [17] in which the output to be predicted is in the form of probabilities of whether a pair of time series should link or not. In practice, mixed models that would both integrate knowledge from a base metric and link constraints will also be considered to allow for a smoother transition between the unsupervised case and the weakly supervised one.

Also, we will push an effort towards metric learning for the class of alignment-based metrics for time series. Here, we will extend the framework of [18] that learns a metric between sequence elements (called feature space metric as opposed to the metric operating at the time series level) based on DTW alignments. Indeed, it is expected that the smoothness of soft-alignment metrics [4] will ease the learning of an adequate feature space metric. Also, as opposed to [18], we will not ground the learning of the feature space metric on ground-truth alignments but we will rather consider alignments as self-supervised information and have the model push matching features closer in feature space. Based on previous work by MATS members [43], this work could also cover automatic hyper-parameter setting.

*Task 3.3: Domain adaptation*

*People involved:* Laetitia Chapel, Nicolas Courty, Titouan Vayer, Romain Tavenard

When little supervised information is available, being able to rely on additional supervised information related to data acquired in related but slightly different context is highly desirable. For example, in a land cover classification setting, being able to re-use data acquired a previous year (and their associated labels) in order to generate new, up-to-date, maps without resorting to a new exhaustive field campaign is of high interest for end-users [21].

We will hence design domain adaptation strategies able learn the temporal transformation needed to embed time series from one domain into another. Based on previous works of members of the MATS team [44, 45], we will consider the optimal transport framework as our main tool to tackle this domain adaptation problem. In order to take into account the temporal nature of our data, two research directions will be explored. First, we will aim at integrating the information carried out by the temporal structure directly in the OT problem. In particular, the lead of defining a dedicated regularization term shall be explored. Second, we will consider the integration of the structure directly inside the distance matrix between the data, building upon the notion of Gromov-Wasserstein distances for instance [46].

**Task 4: Application to environment**

In this project, we will cover three application settings that are detailed in tasks 4.1 to 4.3 below. These application settings are related to key environmental issues for which we believe the methods developed in the MATS project should help decision-making. All three applications are supported by dedicated datasets made available to the MATS research team through Kalideos project and *Zone Atelier Armorique* (*cf.* paragraph "Data means" of Section 2.2.2 for more details).

*Task 4.1: Agricultural practices*

*People involved:* Emilien Alvarez, Anthony Bagnall, Jason Lines, François Petitjean, Romain Tavenard

Agricultural practices have a huge impact on domains as diverse as food safety [47], biodiversity [48] and environmental sustainability in general. Providing decision makers with efficient tools for the identification of such practices is hence of prime importance.

In this task, we will build upon the methods developed in the project to produce large-scale yet precise land cover maps. First, time series classification methods developed in task 1, paired with spatial techniques emerging from task 2.3, will be used to produce object-level classifications from high temporal resolution data. Such classification will be further robustified by considering land cover succession patterns for agricultural parcels [49]. Knowing that such patterns exist, we will use the probabilistic output of our time series classifiers and mix them with (learned) transition probabilities to get robustified class estimations. Moreover, knowledge about these patterns will be a way to generate more realistic evolution scenarios when coupled with generative models from task 2. Finally, succession patterns will be a great by-product by themselves as they are of prime interest for water quality monitoring [50] or biodiversity studies [48].

However, one has to take into account the high human cost associated to the generation of reliable annotations in this context. Indeed, these annotations either emerge from field campaigns or as the output of a tedious photo-interpretation process by experts of the field. This indicates a high interest for application of the domain adaptation methods issued from task 3.3 in that context. Also, semi-supervised approaches (*cf.* task 3.1) that would leverage on the vast amount of available unlabelled data and make the most out of the few labelled samples would be of great help for this task.

Finally, a focus will be given to specific areas in which standard supervised classification settings would face limitations. Indeed, in some cases, a given land cover class can derive in so many different profiles depending on the meteorological conditions that classes themselves are ill-defined and using supervised classification is unlikely to help. For example, in the case of marshlands, a grassland could have completely different reflectance values depending on whether it is flooded or not, and unsupervised approaches that would allow end-users to browse the data collection should be preferred to supervised alternatives.

*Task 4.2: Forest preservation*

*People involved:* Laetitia Chapel, Pierre Gloaguen, Johan Oszwald, Romain Tavenard

United Nations consider forests as a “critical part of the solution to the unprecedented demand for water, food and energy”. In this context, we will provide methods to quickly identify the presence of various degradation / regeneration patterns from satellite image time series, as well as anomaly detection tools to identify anomalous tree growth in large datasets.

A first important issue we will be interested in is that of forest dieback. It is related to tree diseases and a better understanding of this phenomenon could help set up efficient forest management strategies to avoid such diseases. This issue is especially challenging in terms of machine learning since vegetation index profiles are likely to differ between tree species and across years and we believe anomaly detection algorithms developed in task 2 could be beneficial here. In this context, online detection tools that would be able to raise alerts would be especially valuable for forest managers to push forward investigations in suspicious areas.

Another related application context is that of forest deterioration monitoring. In this context, one would like to assess for the impact of human activity on the conservation of forests. Here, the idea would be to identify several types of deterioration and temporally locate them. This could be done through pattern identification and localization, using methods from tasks 1 and 2.

*Task 4.3: Air quality monitoring*

*People involved:* Thomas Corpetti, Romain Tavenard

Existing computational fluid dynamics (CFD) models for pollutant dispersion are at the moment very efficient but their use in operational forecasting systems is still limited since they require solving high order Navier-Stokes equations which are delicate in practice, especially in urban environments where border conditions and obstacles (streets, buildings, ...) are tricky to take into account.

In this application, we plan to take benefit of the numerous existing observations of winds (issued from meteorological forecasts or directly computed from satellite images to recover fine scale turbulence [51]) to design key spatio-temporal patterns using spatio-temporal GAN models from task 2.3. These patterns will be used for two applications. First, by linking them to existing air quality indexes (very good, good, average, poor, very poor) measured from on-site sensors, one can easily forecast air quality from existing winds by identifying/classifying the various spatio-temporal patterns. Second, these spatio-temporal patterns can be used as basis functions to simulate wind evolution that can be used as inputs for pollution dispersion models (in practice we will use the one proposed in [52]).

**List of project deliverables**

Based on the previous description of the organization of the project by task, we derive a list of deliverables that will be needed to ensure the smooth running of the project. These deliverables are in line with the Gantt project presented in Figure 1. Note that all deliverables listed for tasks 1 to 3 concern both model definitions and their actual implementations.

| Deliverable description | | Due by | Produced by | Required for |
|---|---|---|---|---|
| D0.a | Project website | T0+6 | T0.1 | – |
| D0.b | Consortium agreement | T0+6 | T0.1 | – |
| D0.c | Organization of a workshop on machine learning for time series | T0+48 | T0.1 | – |
| D1.a | Time series classification model | T0+18 | T1.1, T1.2 | T4.1 |
| D2.1.a | Temporal normality model & online detection strategy | T0+36 | T2.1 | T4.2 |
| D2.2.a | Temporal GAN model | T0+24 | T2.2 | T2.3, T3.1, T4.1 |
| D2.3.a | Spatio-temporal model for pattern extraction | T0+36 | T2.3 | T4.3 |
| D3.1.a | Semi-supervised model for time series | T0+36 | T3.1 | T4.1 |
| D3.3.a | Domain adaptation strategy using optimal transport | T0+24 | T3.3 | T4.1 |
| D4.1.a | Annual land cover maps for the ZAA region | T0+48 | T4.1 | – |
| D4.2.a | Potential dieback maps for a forest located in the ZAA area | T0+48 | T4.2 | – |
| D4.2.b | Forest degradation patterns for a forest endangered by human activity | T0+48 | T4.2 | – |
| D4.3.a | Spatio-temporal patterns for pollution dispersion | T0+48 | T4.3 | – |

### 2.2.2 Available and requested means

**Human means**

As described in more details in Section 2.1, MATS project gathers researchers with complementary backgrounds from time series analysis, machine learning/statistics and remote sensing fields. Overall, implication of non-funded members of the MATS project will cover a total of 119 persons.month, while funding is requested for 48 persons.month. Among members of the project that will not be funded by the project, a PhD student and a post-doctoral researcher are funded through research project related to MATS to some extent. First, in the context of the SESAME project (ANR ASTRID call)[8], Pierre Gloaguen is interested in the development of temporal topic models for maritime surveillance applications. Second, Titouan Vayer has started his PhD work this year under my supervision and that of Prof. Laetitia Chapel and Nicolas Courty, with a PhD scholarship from Université de Bretagne Sud and connexions with the ANR-funded OATMIL project.[9] Titouan's work is concerned with the design of optimal transport tools for structured data and time is one kind of structure concerned by his work.

Requested funding corresponds to one PhD scholarship as well as one year of engineer work.

The recruited PhD student will be supervised by Romain Tavenard and Thomas Corpetti. The applicant should have a strong level in applied mathematics, machine learning and scientific programming. We will be preferably seeking first rate students with specializations in applied mathematics or computer science. The PhD student will tackle most of the problems related to time series classification (T1.1 and T1.2) and semi-supervised learning (T3.1).

We will also hire an engineer with a solid background in data science and experience in scientific programming to work on the integration of MATS methods in `tslearn`. To keep the `tslearn` toolbox appealing to a wide audience outside the MATS project itself, (s)he will also work on the implementation of standard baseline methods.

[8] http://www.agence-nationale-recherche.fr/Project-ANR-16-ASTR-0026
[9] http://www.agence-nationale-recherche.fr/Project-ANR-17-CE23-0012

**Data means**

Three application settings will be tackled in the context of this project. For them to be successful, we have already identified datasets on which we could rely to assess the performance of our algorithms. These datasets are available through projects in which the LETG partner is involved. First, Kalideos[10] project has started in 2001 and aims at building applications on top of high-quality pre-processed CNES image products (CNES being the french space agency). In Kalideos, five study sites are active, among which the Brittany site develops methodological approaches for a remote-sensing-based monitoring of land cover evolutions. Second, *Zone Atelier Armorique* (ZAA)[11] aims at better understanding existing relationships between political policies, land use and ecological processes. It has been set up between 1993 and 2000 and covers 130 km$^2$. Note that study sites for these projects overlap such that ground truth gathered through ZAA could be used in conjunction with Kalideos data. This study site features agricultural practice issues as well as forest preservation challenges and pollution monitoring concerns in urban areas.

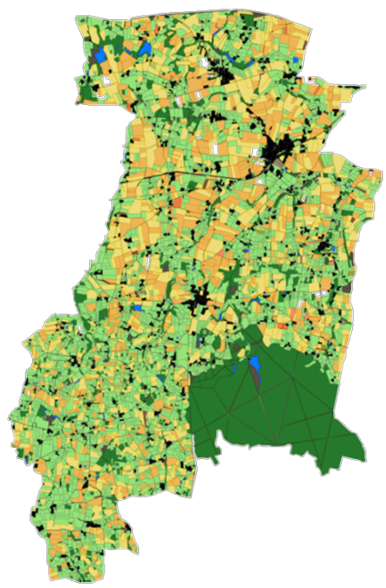

Figure 2: Example land cover map for the *Zone Atelier Armorique* area featuring 12 agricultural classes.

Through these projects, we will for example have access to land cover information for the ZAA area over the last 20 years (see Figure 2 for an example land cover map), which perfectly fits the needs of task 4.1. Other high quality datasets that match the needs of tasks 4.2 (focused on the *forêt de Villecartier*) and 4.3 (using the data arising from city of Rennes' sensor network) will be used to showcase the benefits of our methods in real-life settings.

**Other means**

As stated above, one important mean in terms of dissemination will be the `tslearn` toolbox.[12] `tslearn` is an open-source Python toolbox created by MATS coordinator in 2017. It relies on `scikit-learn` and implements various machine learning methods for time series. It has attracted attention from a large panel of researchers and practitioners (200+ stargazers on GitHub in less than a year) with backgrounds as diverse as machine learning, finance, astronomy, among others. `tslearn` is, at the time of the proposal, made of 2,000+ lines of Python code with high standards in terms of code quality (including automatic build tests, unit tests with a code coverage level over 90% and exhaustive documentation). Methods developed in the context of MATS will be implemented in `tslearn` so as to (i) ease their use in different application contexts and (ii) favor the development of novel methodological works based on these methods.

Other requested means cover the organization of a workshop on machine learning for time series for a better dissemination and visibility of the project. The workshop will either be part of a major conference or an independent workshop. Funding will be used to invite international experts in the domain and will help foster new collaborations.

Funding is also requested for computers and travel expenses. Targeted computers are powerful stations equipped for scientific computing. Travel expenses mainly regard working meeting among MATS members and participation to international conferences. Finally, 8% of the total amount is retained for management and

[10] https://www.kalideos.fr/
[11] https://osur.univ-rennes1.fr/za-armorique/
[12] https://github.com/rtavenar/tslearn

administrative fees.

**Summary of requested means by item of expenditure**

| | |
|---|---|
| **Staff expenses** | |
| 1 PhD student (36 months) | 100 k€ |
| 1 engineer (12 months) | 40 k€ |
| **Sub-total for staff expenses** | **140 k€** |
| **Travel expenses** | |
| Organization of a workshop | 10 k€ |
| Project meetings (2 per year) | 8 k€ |
| Participation to conferences (3 per year) | 22 k€ |
| **Sub-total for travel expenses** | **40 k€** |
| **Other expenses** | |
| Stations / laptops for software development (4 computers) | 10 k€ |
| Publication fees | 5 k€ |
| Consortium agreement (SATT) | 4 k€ |
| Administrative fees (8 %) | 16 k€ |
| **Sub-total for other expenses** | **36 k€** |
| **Sub-total** | **215 k€** |
| **Requested funding** | **215 k€** |

## 2.3 Risk assessment

The MATS project implies risks ranging from moderate to important. Tasks 1 and 3 will benefit from previous works of the project members and its main research directions are suggested by promising preliminary results. In more details, tasks 1.1 and 1.2 are extensions of seminal works of the team [5, 12, 28] and can be considered as of low risk. Similarly, task 2.1 consists in the design of a continuous-time extension of [26], while task 3.2 will build upon the LDPS model [17], also introduced by members of the team.

Tasks 2.2 and 3.1 share similar risk as they both rely on the development of novel temporal GAN models. On this point, we will consider variants associated with very different levels of risk so as to ensure at least one such model (*eg.* convolutional GAN for regularly sampled, fixed sized, time series) can be proposed in the frame of the project while not preventing ourselves from working on more elaborate scenarios associated with high potential impact. Task 2.3 includes elements of higher risk (spatio-temporal models represent a new research direction for MATS members) for which existing less risky alternatives could be considered at some point if needed. As an example, building on models developed in other tasks, spatial coherence could be enforced in an *a posteriori* step if we do not succeed in proposing fully integrated solutions. Though task 3.3 relies on previous works of the team [44], the development of temporal-structure-aware optimal transport is still challenging and less challenging alternatives (*eg.* kernel-based methods) could be considered if necessary. However, this high risk related to tasks 2 and 3 also suggests high pay-off if successful.

Success of the application tasks (T4.*), will rely on the availability of related methodological developments, their implementation in the `tslearn` toolbox and the access to high-quality datasets. If these three conditions are met, the risk related to task 4 is low, and we have shown in Section 2.2.2 that the last two conditions are guaranteed by already available means. If some methodological developments were to be lacking, we would ground our applicative developments on technical methodologies that are easy to acquire (T1.1, 1.2, 2.1, 3.2) for which specific adaptation will be made to meet the application requirements.

Measuring the success of the project will be made possible by indicators related to publications in international renowned journals or conferences and access statistics of the `tslearn` toolbox. Most publications generated by the project will be in the field of machine learning / data science, though it is expected that publications related to task 4 will be targeted towards the remote sensing audience.

# 3 Impact and benefits of the project

From a scientific point of view, the MATS project is expected to provide a substantial breakthrough in the field of machine learning for time series. This project will notably feature ground-breaking propositions for time-series-specific methods in diverse machine learning problems, which are currently missing in the state-of-the-art. Moreover, all developments will be included in the `tslearn` open source toolbox. The latter toolbox is developed by the project coordinator and already features a wide range of state-of-the-art time-series specific machine learning algorithms. Finally, application of these novel methods to help face crucial environmental issues will be directly addressed in the context of the project. Three key application settings will be considered. First, to deal with the impact of agricultural practices on the environment, large-scale yet precise land cover maps will be produced. Second, forest preservation issues will be tackled through anomaly detection and motif extraction approaches. Finally, air pollution monitoring will benefit from spatio-temporal models developed in MATS for a better understanding of fluid motion patterns involved.

Diffusion of the outputs of the MATS project will be done through publications in major international conferences and journals of the fields of machine learning (NIPS, ICML, ECML, JMLR, *etc.*) and remote sensing (IEEE TGRS, IEEE JSTARS, Multitemp, *etc.*). Also, a Python toolbox (`tslearn`) will be used to favor a wide adoption of the tools developed in the context of the project. Moreover, an international workshop on the topics of the MATS project will be organized to participate in the dissemination of the project outputs and foster future collaborations of the project participants with top-level international researchers. Finally, methodological developments produced in the context of the project will be used to feed academic teaching programs, as well as to propose new actions in french *Groupements de Recherche* (typically for the GDR MaDICS[13] whose research interests cover MATS topics).

MATS fits into challenge 7 “Société de l’information et de la communication” of the ANR action plan for the 2018 project call and more precisely to axis 4 “Données, Connaissances, Big Data, Contenus multimédias – Intelligence Artificielle” and the related “France is AI” incentive.[14] Indeed, it aims at developing ground-breaking machine learning algorithms for time series data, with a focus on scalable algorithms. As such, it fits in Orientation 28 of Stratégie Nationale de Recherche (SNR, défi 7) that concerns “Exploitation des grandes masses de données”. It is also related, through its application task, to challenge 1 “Gestion sobre des ressources et adaptation au changement climatique” and its Orientations 1 (“Suivi intelligent du système terre”) and 2 (“Gestion durable des ressources naturelles”).

Finally we believe that this project is a perfect fit for a young researcher project (JCJC) as it will, through the recruitment of a PhD student and an engineer, help the coordinator push forward his research interests in time series analysis and machine learning in the applicative context of a geography lab such as LETG. This project will also be an opportunity to develop fruitful international connections with renowned researchers through both collaborations in MATS tasks and the organization of a workshop. The MATS project is hence expected to provide opportunities for the project coordinator to become a reference in the field of machine learning for time series both at the national and international scales. Upon success, it will pave the way for an ERC application of Romain Tavenard on this topic.

[13] http://www.madics.fr
[14] https://franceisai.com

## 4 References related to the project

[1] M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley and X. Huang. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote sensing of Environment* 114.1 (2010), pp. 168–182.

[2] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing* 26.1 (1978), pp. 43–49.

[3] M. Cuturi. Fast global alignment kernels. In: *Proceedings of the International Conference on Machine Learning*. 2011, pp. 929–936.

[4] M. Cuturi and M. Blondel. Soft-DTW: a Differentiable Loss Function for Time-Series. In: *Proceedings of the International Conference on Machine Learning*. PMLR, 2017, pp. 894–903.

[5] A. Bailly, S. Malinowski, R. Tavenard, L. Chapel and T. Guyet. Dense Bag-of-Temporal-SIFT-Words for Time Series Classification. In: *Advanced Analysis and Learning on Temporal Data*. Springer, 2016.

[6] P. Schäfer. The BOSS is concerned with time series classification in the presence of noise. *Data Mining and Knowledge Discovery* 29.6 (2015), pp. 1505–1530.

[7] L. Ye and E. Keogh. Time series shapelets: a new primitive for data mining. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data mining*. 2009, pp. 947–956.

[8] I. Karlsson, P. Papapetrou and H. Boström. Generalized random shapelet forests. *Data Mining and Knowledge Discovery* 30.5 (2016), pp. 1053–1085.

[9] J. Grabocka, N. Schilling, M. Wistuba and L. Schmidt-Thieme. Learning time-series shapelets. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data mining*. 2014, pp. 392–401.

[10] Z. Wang, W. Yan and T. Oates. Time series classification from scratch with deep neural networks: A strong baseline. In: *Proceedings of the International Joint Conference on Neural Networks*. 2017, pp. 1578–1585.

[11] Y. Zheng, Q. Liu, E. Chen, Y. Ge and J. L. Zhao. Time series classification using multi-channels deep convolutional neural networks. In: *Proceedings of the International Conference on Web-Age Information Management*. 2014, pp. 298–310.

[12] A. Le Guennec, S. Malinowski and R. Tavenard. Data Augmentation for Time Series Classification using Convolutional Neural Networks. In: *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*. 2016.

[13] M. Hüsken and P. Stagge. Recurrent neural networks for time series classification. *Neurocomputing* 50 (2003), pp. 223–235.

[14] F. Petitjean, A. Ketterlin and P. Gançarski. A global averaging method for dynamic time warping, with applications to clustering. *Elsevier Pattern Recognition* 44.3 (2011), pp. 678–693.

[15] J. Paparrizos and L. Gravano. k-shape: Efficient and accurate clustering of time series. In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*. 2015, pp. 1855–1870.

[16] J. Zakaria, A. Mueen and E. Keogh. Clustering time series using unsupervised-shapelets. In: *Proceedings of the IEEE International Conference on Data Mining*. 2012, pp. 785–794.

[17] A. Lods, S. Malinowski, R. Tavenard and L. Amsaleg. Learning DTW-Preserving Shapelets. In: *International Symposium on Intelligent Data Analysis*. 2017, pp. 198–209.

[18] D. Garreau, R. Lajugie, S. Arlot and F. Bach. Metric learning for temporal sequence alignment. In: *Advances in Neural Information Processing Systems*. 2014, pp. 1817–1825.

[19] L. Wei and E. Keogh. Semi-supervised time series classification. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data mining*. 2006, pp. 748–753.

[20] M. González, C. Bergmeir, I. Triguero, Y. Rodríguez and J. M. Benítez. Self-labeling techniques for semi-supervised time series classification: an empirical study. *Knowledge and Information Systems* (2017), pp. 1–36.

[21] A. Bailly, L. Chapel, R. Tavenard and G. Camps-Valls. Nonlinear Time-Series Adaptation for Land Cover Classification. *IEEE Geoscience and Remote Sensing Letters* (2017), pp. 896–900.

[22] E. Gurarie, C. H. Fleming, W. F. Fagan, K. L. Laidre, J. Hernández-Pliego and O. Ovaskainen. Correlated velocity models as a fundamental unit of animal movement: synthesis and applications. *Movement Ecology* 5.1 (2017), p. 13.

[23] G. Walker. On Periodicity in Series of Related Terms. *Proceedings of the Royal Society of London. Series A.* 131.818 (1931), pp. 518–532.

[24] J. Contreras, R. Espinola, F. J. Nogales and A. J. Conejo. ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems* 18.3 (2003), pp. 1014–1020.

[25] R. Emonet, J. Varadarajan and J. M. Odobez. Temporal Analysis of Motif Mixtures Using Dirichlet Processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.1 (2014), pp. 140–156.

[26] R. Tavenard, R. Emonet and J.-M. Odobez. Time-Sensitive Topic Models for Action Recognition in Videos. In: *Proceedings of the IEEE International Conference on Image Processing*. 2013.

[27] J. Lines, L. M. Davis, J. Hills and A. Bagnall. A shapelet transform for time series classification. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data mining*. 2012, pp. 289–297.

[28] R. Tavenard, S. Malinowski, L. Chapel, A. Bailly, H. Sanchez and B. Bustos. Efficient Temporal Kernels between Feature Sets for Time Series Classification. In: *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery*. 2017, pp. 528–543.

[29] N. Simon, J. Friedman, T. Hastie and R. Tibshirani. A sparse-group lasso. *Journal of Computational and Graphical Statistics* 22.2 (2013), pp. 231–245.

[30] L. Wu, I. E.-H. Yen, F. Xu, P. Ravikuma and M. Witbrock. D2KE: From Distance to Kernel and Embedding. *arXiv preprint arXiv:1802.04956* (2018).

[31] M. Shah, J. Grabocka, N. Schilling, M. Wistuba and L. Schmidt-Thieme. Learning DTW-shapelets for time-series classification. In: *Proceedings of the ACM International Conference on Data Science*. 2016, p. 3.

[32] I. J. Goodfellow, J. Shlens and C. Szegedy. Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572* (2014).

[33] N. Papernot, P. McDaniel, X. Wu, S. Jha and A. Swami. Distillation as a defense to adversarial perturbations against deep neural networks. In: *Proceedings of the IEEE Symposium on Security and Privacy*. 2016, pp. 582–597.

[34] R. Tavenard and S. Malinowski. Cost-Aware Early Classification of Time Series. In: *Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery*. 2016, pp. 632–647.

[35] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio. Generative adversarial nets. In: *Advances in Neural Information Processing Systems*. 2014, pp. 2672–2680.

[36] M. Saito, E. Matsumoto and S. Saito. Temporal generative adversarial nets with singular value clipping. In: *Proceedings of the International Conference on Computer Vision*. 2017, pp. 2830–2839.

[37] A. Radford, L. Metz and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434* (2015).

[38] D. Britz, A. Goldie, T. Luong and Q. Le. Massive Exploration of Neural Machine Translation Architectures. *arXiv preprint arXiv:1703.03906* (2017).

[39] M. Arjovsky, S. Chintala and L. Bottou. Wasserstein GAN. *arXiv preprint arXiv:1701.07875* (2017).

[40] S. S. Ravindran. A reduced-order approach for optimal control of fluids using proper orthogonal decomposition. *International journal for numerical methods in fluids* 34.5 (2000), pp. 425–448.

[41] S. Chopra, R. Hadsell and Y. LeCun. Learning a similarity metric discriminatively, with application to face verification. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Vol. 1. 2005, pp. 539–546.

[42] E. Hoffer and N. Ailon. Deep metric learning using triplet network. In: *Proceedings of the International Workshop on Similarity-Based Pattern Recognition*. 2015, pp. 84–92.

[43] C. Tan, M. Herrmann, G. Forestier, G. Webb and F. Petitjean. Efficient search of the best warping window for Dynamic Time Warping. In: *Proceedings of the SIAM International Conference on Data Mining*. 2018.

[44] N. Courty, R. Flamary, D. Tuia and A. Rakotomamonjy. Optimal transport for domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.9 (2017), pp. 1853–1865.

[45] N. Courty, R. Flamary, A. Habrard and A. Rakotomamonjy. Joint distribution optimal transportation for domain adaptation. In: *Advances in Neural Information Processing Systems*. 2017, pp. 3733–3742.

[46] G. Peyré, M. Cuturi and J. Solomon. Gromov-Wasserstein Averaging of Kernel and Distance Matrices. In: *Proceedings of the International Conference on Machine Learning*. 2016, pp. 2664–2672.

[47] D. Bourn and J. Prescott. A comparison of the nutritional value, sensory qualities, and food safety of organically and conventionally produced foods. *Critical reviews in food science and nutrition* 42.1 (2002), pp. 1–34.

[48] A. McLaughlin and P. Mineau. The impact of agricultural practices on biodiversity. *Agriculture, Ecosystems & Environment* 55.3 (1995), pp. 201–212.

[49] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin and I. Rodes. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sensing* 9.1 (2017).

[50] A. Aubert, R. Tavenard, R. Emonet, A. De Lavenne, S. Malinowski, T. Guyet, R. Quiniou, J.-M. Odobez, P. Mérot and C. Gascuel-Odoux. Clustering Flood Events from Water Quality Time-Series using Latent Dirichlet Allocation Model. *Water Resources Research* 49.12 (2013), pp. 8187–8199.

[51] P. Zille, T. Corpetti, L. Shao and C. Xu. Super-resolution of turbulent passive scalar images using data assimilation. *Experiments in Fluids* 57.2 (2016), p. 21.

[52] Y. Liu, G. Cui, Z. Wang and Z. Zhang. Large eddy simulation of wind field and pollutant dispersion in downtown Macao. *Atmospheric environment* 45.17 (2011), pp. 2849–2859.