Estimating the Dominant Frequencies of Real-Valued Cyclic Processes in Natural Environments for Long-Term Operation of Autonomous Robots

Tomáš Vintr[1,3], Jan Blaha[1], Jiří Ulrich[1], Tomáš Rouček[1], George Broughton[1], Tom Duckett[2], Farshad Arvin[3], and Tomáš Krajník[1]

***Abstract*— Autonomous mobile robot localisation, planning, and navigation methods typically rely on environmental representations. Previous research has shown that the temporal dynamics captured by specialized models help autonomous robots to operate for longer periods with better efficiency. One of the leading approaches uses maps enhanced by frequency analysis to model and predict repeating cycles of activity in the environment. However, this approach implicitly relies on prior knowledge of the expected periodicities, such as days and weeks, encoded by the human designers of the system. This paper presents a new method to automatically search the robot's observations for dominant frequencies, leading to a more general method than the previous approach for frequency map enhancement. The proposed algorithm extends the problem definition from binary observations also to real-valued time series, making it applicable to a broader spectrum of robotic tasks. We show that the new method can be implemented in robotic systems operating without prior knowledge of the underlying processes that influence the dynamics of the working environment across a wide variety of tasks similar to the long-standing state-of-the-art. We hypothesise that an autonomous robot using the proposed improvement can be deployed to unprecedented environments.**

Keywords: **long-term autonomy, spectral decomposition, spatio-temporal modelling**

I. Introduction

The rise of autonomous mobile robotics comes together with the development of their ability to understand the surrounding world. More recent work has shown that the maps and models representing the environment benefit from the inclusion of temporal dynamics and the capability of these models to predict future environment states. This has brought questions of how to best explore, capture, represent, and exploit dynamics, suppress negligible trends, and discover periodic patterns and rare cyclical events. The topic is complicated by the fact that, while these robots are autonomous, they are operating in the real world and, therefore, are subject to real-world technical limitations. For example, battery charging downtime can create gaps and non-uniformities in observations. Nevertheless, it has been shown that maps that model and forecast environment states allow mobile robots to handle the changes better compared to methods that simply adapt the maps to the changes observed [1], [2], [3]. Maps that explicitly model environment dynamics are called Maps of Dynamics (MoD), which are postulated as "*queryable models of spatial or spatio-temporal patterns of dynamics*" [4], [5].

Generally, we are interested in the class of "Temporal MoDs" as classified in [4]. The Temporal MoD is an environment representation parameterised by time [6]—a spatio-temporal model, usually composed of multiple spatial elements with independent temporal characteristics, representing different dynamics at different locations [7]. Temporal MoDs assume that the patterns of dynamics are time-dependent and repetitive; therefore, the state of the environment in the future can be forecasted. In this paper, we focus specifically and only on individual time series gathered from robotic observation and their automated analysis. It is characteristic that the observations made in time from a moving robot lead to irregular time series with large gaps. Such an irregularity is not the standard case in other time-series analysis methods, e.g., video analysis [8] but it is inevitable for the practical deployment of autonomous mobile robots.

Research into these methods has shown that the primary drivers of changes in natural environments that robots need to deal with tend to be periodic, whereas trends seem to be negligible in the midterm. As a result, one of the most successful ways to model them in a Temporal MoD is to use the approach known as *Frequency Map Enhancement* (FreMEn) [9] derived from the Non-Uniform Fourier Transform (NUFT) [10], [11]. It was demonstrated in [12] that unlike for a reconstruction task where more frequency components lead to better performance, for the forecasting task, one needs to specify a small set of accurately chosen dominant frequencies in order to generalize well. The most important characteristic of the FreMEn method is its ability to gener-

[1] Laboratory of Chronorobotics, Artificial Intelligence Center, Department of Computer Science, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czechia. {tomas.vintr, jan.blaha, jiri.ulrich, tomas.roucek, george.broughton, tomas.krajnik}@fel.cvut.cz

[2]Lincoln Centre for Autonomous Systems. tom.d.duckett@gmail.com

[3]Swarm & Computation Intelligence Laboratory, Department of Computer Science, Durham University, Durham, United Kingdoms. {tomas.vintr, farshad.arvin}@durham.ac.uk

This work was supported by EU project RoboRoyale number 964492 and EU project SensorBees number 101130325, by the Grant Agency of the Czech Technical University in Prague, grant No. SGS22/168/OHK3/3T/13 and project Robotics and Advanced Industrial Production no. CZ.02.01.01/00/22_008/0004590. Special thanks go to Jiří Ulrich and Grzegorz Cielniak. Jiří Ulrich helped open Pandora's box and solved the problems that compelled others to flee. Grzegorz Cielniak sacrificed his privacy and comfort for two years while creating the MHT lecturer office dataset.

alize, allowing for useful short- and mid-term forecasts.

FreMEn introduced periodical dynamics into many discrete robotic representations of an environment, showing that its ability to predict the future structure of an environment improves robotic mapping [13], [9], [14], [15], localisation [12], [16], [9], [17], path planning [18], [19], [20], [21], navigation [22], [23], exploration [24], [25], [26], [27], [28], task scheduling [29], [19], [20], patrolling [24], [30], searching [31], novelty detection [13], [9], [32], [33], activity recognition [34], human-robot interaction [35], [36], [20], [37], and demand forecasting [38], and helped people avoid crowded places during the COVID19 outbreak in Czech Republic [39], [40]. After a while, new robotic methods started to successfully adopt the part of FreMEn that models individual time series as a 'black-box' predictor, for example [14], [17], [37], [21], [23]. However, all those applications modeled phenomena with a sole underlying generative process present in human working environments, especially offices.

In these environments, the underlying dynamics typically form week- and day-long periodic patterns of humans' working routines. The day-based routines also naturally match the day-night changes of an environment driven by illumination. People's routines are further strengthened by the fact that people follow calendars and clocks. Therefore, the set of frequencies analyzed by the original FreMEn was tailored for week-to-hours-long routines and evolved into an integral part of the algorithm [41]. As a result, FreMEn-based methods improved the performance of robots in the previously studied tasks, but their success obscured the fact that the algorithm is optimized exclusively for a specific type of environment. The deployment of the algorithm in environments with general dynamics would require a qualified guess of a suitable set of frequencies and their manual implementation [42].

The problem of specifying a set of candidate angular frequencies is extremely restrictive in the long-term autonomy setting, where the design-time assumptions cannot be guaranteed. We propose an approach in which the most crucial part of NUFT-based Temporal MoDs does not depend on a manually pre-selected set of candidate frequencies. The dominant frequencies are inferred directly from the data using an iterative sampling-based search for a frequency dominant with respect to the other frequencies. This gives a robot the ability to learn temporal patterns without relying on prior knowledge or assumptions about its environment from the designer. We believe it opens the way for pertinent applications of Temporal MoDs in environments where underlying processes are unknown or under research, such as deep sea, outer space, and local ecosystems [43].

II. Related Work

Many robotic methods assume that environmental uncertainty comes primarily from the imperfection of the sensory input [6] and ignore the fact that the environment itself can change over time [44]. Some methods consider environmental changes that significantly influence the deployment of a robot. For example, illumination changes during the day can be mitigated by filtering the received data in the photometric domain [45], [46] or by utilising environment structure [47]. Some changes in the scenes can be modelled using Bayesian probability [48], and seasonal changes can be suppressed by sophisticated forecasting of visual appearance [49], [50].

Philosophically different approaches to environmental changes emerged from the idea that localisation and mapping methods used by autonomous robots need not focus only on accuracy but, more importantly, on their flexibility, robustness and adaptability [51], [3]. The idea evolved during the STRANDS project [52], focusing on the long-term deployment of autonomous robots in natural, human-populated areas, and then expanded during the project STROLL [53]. The STRANDS consortium argued that environmental changes should not be understood as an inevitable source of the gradual degradation of well-established static models (occupancy grids, visual landmarks, edges in topological maps). Instead the natural dynamics of the environment can learned, and exploited by long-term running autonomous systems. The STRANDS team observed that a large part of the dynamics in human environments relates to natural cycles, and they determined that the uncertainty in the models can be represented as a function of time. Through identification of these routines, autonomous robots can learn from and adapt to the changes rather than trying to neglect or suppress them.

A. Non-FreMEn Approaches to Spatio-Temporal Mapping

The most common way to model periodical events in the spatio-temporal maps is to include a priori known periods derived from the designers' expertise in the model's architecture. The straightforward way is to create seasonal windows that include time-specific maps [54], [55], [56], [42], [38], [57]. Another approach is based on kernel warping. The map includes a continuous Gaussian mixture model with predefined periodical kernels [58], [59], [60], [61], [62]. There also exist approaches of manual preprocessing the data to find dominant periods using auto-correlation [63], [64], [65] and then applying the found periods into the architecture of the spatio-temporal maps.

Automatic estimation of periods in the data was targeted in [66] by applying Gaussian processes to air flow data. The model included a covariance matrix consisting of separately calculated spatial and temporal components. Temporal components included temporal decay and periodic components estimated from the data iteratively using frequency analysis. The same group then expanded the idea by proposing Hilbert maps [67], their incremental update [68], Fourier Feature Approximations for Periodic Kernels [69] and its multi-dimensional variant [70]. However, the mathematical and computational complexity limits its usage in autonomous robot systems.

B. Fourier Transform in Long-Term Autonomy

In [13], the authors apply the Fast Fourier Transform (FFT) to the past observed binary states of an occupancy

grid and define the occupancy states as a periodic function of time. *Spectral Mapping* thus acquired the ability to predict the states in the map. Compared to the classical occupancy grid, Spectral Mapping lowered the prediction error by 60%. They optimised the occupancy map by combining Spectral Mapping with an octree-based spatial model [71] into a 4-dimensional model of the environment, FROctomap [72], and integrated this within the Robotic Operating System (ROS). Spectral Mapping was also applied to the topological localisation task [12]. The robot visited 8 different places every 10 minute for one week and captured the visual appearance of its surroundings at every place. The images were processed into image features, and the FFT then modelled the visibility of those features. The robot created a map of features specific to a place and time. The experiments proved the ability of the robot to localise itself using the feature map even after 3 months. The authors found a correlation between the prominence of periods used in the model and the persistence of the map.

The Spectral Mapping method proved that temporal prediction gives autonomous robots an advantage, but inherent FFT principles require a stable observation rate without interruption. It is hard to attain such observations outside a laboratory, especially with autonomous mobile robots. Additionally, Spectral Mapping cannot be updated by adding new data from later observations, limiting long-term deployments. Those problems were addressed by *Frequency Map Enhancement* (FreMEn) [73], [9] which is based on the definition of the *Non-Uniform Discrete Fourier Transform* (NUDFT) [74] rather than FFT and allows building and updating models with irregular data.

FreMEn was applied to the edges of a topological map [18] as a part of the Time-Indexed Navigation Markov Decision Processes. During the experiment, the robot gathered data over 2 months and demonstrated a significant improvement in planning the optimal time to navigate through the environment. To prove the dominance of the proposed approach in planning tasks, FreMEn was applied to a robotic search task [31]. The robot modelled the presence of people at different places in 3 different environments. During the experiment, it planned a path through the environment, intending to find a person as quickly as possible. Compared to the strategy based on the static map, which did not consider human habits, the strategies exploiting the model of usual human behaviours decreased the search time by 25% and the number of visited places by 33%. Aiming to model the number of people in certain areas instead of binary states or the probability, the authors of [35] redefine FreMEn, which resulted in an iterative approach denoted as Addition Amplitude Model (AAM). Recently, FreMEn was used to build a map that considered not only the human presence but also their direction [14], [27]. Such a map provides information that allows for planning navigation in a crowd while adhering to human flows [19]. A non-traditional application of FreMEn can be found in [37], where FreMEn served to ensure the optimal division of tasks between multiple robots operating in a human-populated environment.

FreMEn proved its superiority to the classic robotic approaches in mapping for navigation and provided original insights into the dynamics of the environment. The question of whether it is possible to utilise FreMEn's qualities in exploration tasks arose. The initial experiments on life-long spatio-temporal exploration provided a hypothesis that the best time and place to explore correlates with the uncertainty of FreMEn [24]. That led to the development of the *information-based Monte-Carlo scheduler* for spatio-temporal exploration [25]. However, visiting places when the robot is still determining what to expect can be contrary to the robot's aims, for example, avoiding busy human flows. This contradiction, known as the exploration-exploitation dilemma, was thoroughly studied in [29]. The effort was crowned by defining the life-long spatio-temporal exploration of dynamic environments [26]. Various exploration strategies were later studied over the spatio-temporal-directional maps [27].

The success of applying FreMEn into various robotic tasks led to its integration into the STRANDS system [52]. The STRANDS system controlled two robots deployed to 2 different environments with different roles. One of them was a security robot operating in an office building in the UK. The second one operated in an Austrian care home with an overall autonomous deployment of 8 months. As the robot had to fulfil its task while simultaneously building and updating its map, the most promising exploration-exploitation strategy was successfully integrated into the STRANDS system. The gradual growth of its performance in real deployments proved the theoretical conclusions [36]. It should be noted that FreMEn, as a mapping, navigation, and planning system, fulfilled the strict requirements of the privacy policy as it did not need to store any visual images.

C. Closer Look at FreMEn

1) Definition: Generally, FreMEn [24], [9] replaces stationary uncertainty models of binary states in robot maps with functions of time, represented by their frequency spectra. For example, the probability of grid cell occupancy is not modelled by a single probability value, updated only by direct observation. Rather, each occupancy grid cell contains a frequency spectrum-based time series model derived by using the NUDFT over previous observations. As the time series decomposition is not used for reconstruction but for prediction, only a few of the most prominent coefficients of the Fourier spectrum are chosen to serve as parameters of the time series model. The number of these prominent coefficients is referred to as the *order of model*.

The difference between FreMEn's decomposition and the decomposition using NUDFT lies in the ability of FreMEn to incorporate new observations incrementally, thus being able to provide predictions at any time during the robot operation. This feature enables FreMEn's on-the-fly learning from sparse and irregular data [76]. Let us assume a set of candidate angular frequencies $\omega_k \in \Omega$ and the order of the

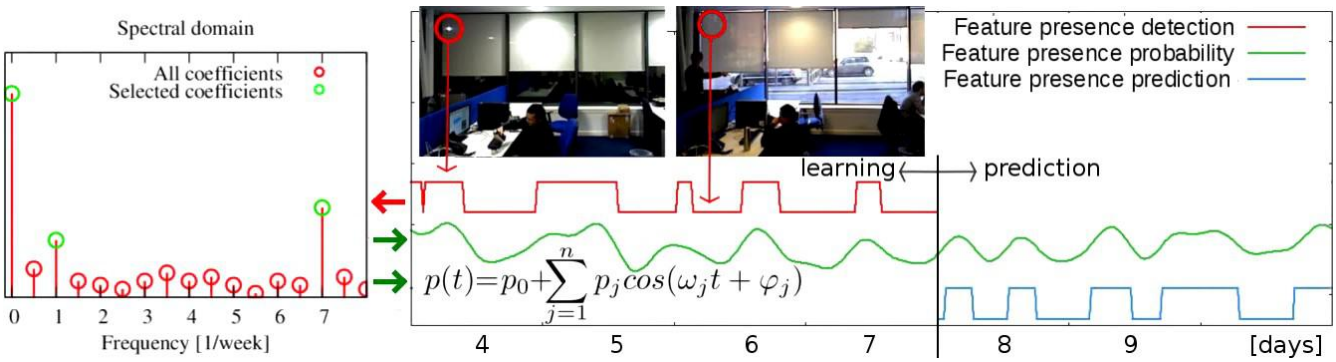

Fig. 1. Temporal MoDs have been demonstrated to bring advantages in long-term autonomy, for example, in tasks such as visual localization. Robotic observations provide us with a series (centre panel, red) of visibility observations of one specific image feature. In this example, it is a reflection of a light source (centre panel, red circles). The time series is decomposed in the spectral domain (left panel), and the dominant frequencies (green) are selected for the model (centre panel, green). The visibility of the feature can then be forecasted (right panel, blue) in the sense of its probability (right panel, green). Courtesy of [75], ©2020, IEEE.

model o. Starting with $n=0$ observations, FreMEn updates the model parameters by adding a new measurement $s(t)$ in the following way [41], [76]:

- update to the mean probability

$$\mu \leftarrow \frac{1}{n+1}\left(n\mu + s(t)\right),$$

- $\forall\omega_k$ update to the state spectrum

$$\alpha_k \leftarrow \frac{1}{n+1}\left(n\alpha_k + s(t)e^{jt\omega_k}\right),$$

- $\forall\omega_k$ update to the observation spectrum

$$\beta_k \leftarrow \frac{1}{n+1}\left(n\beta_k + e^{jt\omega_k}\right),$$

- update to the number of observations

$$n \leftarrow n+1.$$

Then a prediction can be performed at any time t_0 considering all previous measurements as follows:

- $\forall k$ calculate the predictive spectrum $\gamma_k \leftarrow \alpha_k - \mu\beta_k$,
- choose o components $\{\gamma_c\}_1^o$ with the highest $|\gamma_k|$,
- and predict the state $s(t_0)$:

$$s(t_0) = \begin{cases} 1 & \text{if } \mu + 2\sum_1^o |\gamma_c| \cos\left(t_0\omega_c - \arg(\gamma_c)\right) > 0.5 \\ 0 & \text{otherwise.} \end{cases}$$

FreMEn was also defined as an estimator of the probability of occupied state occurrence $P(s(t)=1)$ [25], [9] such that the prediction (II-C.1) was transformed into:

$$P(s(t_0)=1) = \\ = \min\left(\max\left(\mu + 2\sum_1^o |\gamma_c| \cos\left(t_0\omega_c - \arg(\gamma_c)\right), 0\right), 1\right), \quad (1)$$

see an example application that forecasts image feature visibility in Figure 1.

The considered angular frequencies ω_k and the order of the model o need to be chosen wisely in advance. The original advice for the set of candidate angular frequencies was proposed in [25] as $\omega_k \in \Omega_{24} = \{2\pi k/86400\}_{k=1}^{24}$ (note that one day consists of 86400 seconds).

2) Parameters: The set Ω_{24} served well in robotic applications where the most significant dynamics came from day/night changes, like topological localisation [12] based on the visibility of features in the video frames, or models of human behaviour calculated over several days using datasets that do not cover weekly routines [25]. However, it is possible to deduce that the set of candidate angular frequencies used by FreMEn includes the angular frequency $2\pi/604800$, for example, from notes in results [24], from the graphs [18], or by finding a back-reference from one article [25] to another [24]. By the deep study of the associated source code [41], one can uncover that the standard set of candidate angular frequencies in FreMEn and FreMEn-derived methods is $\omega_k \in \Omega_{168} = \{2\pi k/604800\}_{k=1}^{168}$, which was not discussed in any work known to authors. The set $\Omega_{168} \supset \Omega_{24}$ includes the frequencies once-per-week and once-per-day, which reflects the environmental dynamics derived from the fact that the majority of the population work on a daily and weekly basis [17].

In the original paper [9], it is advised to choose the best model order o by testing the prediction error of models with different o trained on the data gathered before the last model update and tested over the data gathered afterwards. Such an approach is justified in the incremental updating of the model by an autonomous robot, whose life cycle consists of working and charging. However, in some experiments, in which the incremental updating has no meaning because all the data are already known, the authors of the experiments omitted the question 'how to divide the training dataset', which is not apparent from the original advice [9], and chose o, for example, arbitrarily [25], [32], or by optimising the models over the testing data [17], [15].

Note that choosing order o as an arbitrary constant can be successful in some experimental settings but goes against the original idea. An important part of FreMEn's applications is its ability to recognize the periodic behaviour of the observed environmental phenomena. That includes choosing a *null complexity* $o=0$ and forecasting only a time-independent probability (constant value) when appropriate.

Deciding on the complexity of the explanatory model is directly responsible for the ability to generalize from the observed dynamics.

3) Shortcomings: Although the non-rigorous approach to estimate the hyperparameter o is not primarily the problem of FreMEn, it accents the weaknesses of FreMEn, which is the necessity to manually choose the correct set of candidate angular frequencies $\Omega_{candidates}$. This set needs to include the angular frequencies expected by the researchers' knowledge about the environment, $\Omega_{expected} \subset \Omega_{candidates}$. The researchers then use FreMEn to calculate the amplitudes of the $\Omega_{candidates}$ and expect that an optimal model order is not much higher than $||\Omega_{expected}||$. Although it is common in the time series analysis to provide a chosen mathematical tool with the known frequencies like once per day, once per week, or four times a year, the original aim of FreMEn, as we understand that, was to equip autonomous robots with the ability to learn the spatio-temporal structure of their environment on their own. In this paper, we primarily focus on developing this original intention of FreMEn by removing the dependence on a manual choice of candidate frequencies Ω.

III. Proposed Approach

A. Main Idea

The FreMEn-related Temporal MoDs shift from neglecting dynamics to anticipating environmental changes by modeling cyclic, stationary patterns of state [77] or domain [78] changes, improving long-term autonomy and robotic operations. However, contrary to expectations, the original FreMEn approach cannot, in general, find the most significant frequencies in the data because the set of frequencies is a necessary and manually chosen input to the algorithm [41]. To target the ability to estimate dominant frequencies that provide generalisation with sufficient forecasting potential, we step back from incremental learning defined in FreMEn to the decomposition of a batch of data based on a non-uniform discrete Fourier transform.

It is possible to find the frequency with the (almost) highest amplitude using, for example, brute force. However, finding multiple frequencies and their correct amplitudes that provide good prediction is problematic due to parasitic frequencies. Inspired by the construction of Warped Hypertime [17] and AAM [35], we propose an iterative version of searching for the most prominent frequencies in the data. Starting from the model order $o = 0$, which is equivalent to averaging of the data, we subtract the model from the data and find the frequency with the highest amplitude in the newly created error time series.

We lighten the exhaustive search for the global maximum by an iterative sampling-based search for a prominent maximum of amplitudes, which allows us to define a termination condition when the search cannot reach the result. On the other hand, the search can get stuck in a local maximum and provide an incorrect model. Therefore we also propose an initialisation stage that consists of multiple model building over several parts of the training data. Those models provide the candidate frequencies to the main search algorithm as starting positions, which lowers the probability of being stuck in a local maximum.

B. Method Overview

We propose a method derived from FreMEn that can be used for the probability of the state prediction and prediction of time series with a negligible trend. The prediction function needs to be selected from (1) and (2) based on the task or demand.

Let us loosen the restriction for time series from binary-only $s(t)$ to a real-valued one $R(t)$. Then we can define the prediction at time t_0 of a model $M^q(t_0)$ with order q as a regression:

$$M^q(t_0) = \gamma_0 + 2\sum_{1}^{q} |\gamma_c| \cos\left(t_0\omega_c - \arg(\gamma_c)\right), \tag{2}$$

where $\gamma_0 = \overline{R(t)}$ is an average of time series values $R(t_i)$, $i = 1 \ldots n$ and n is the number of measurements, and γ_c are coefficients of the Fourier series corresponding to the considered angular frequencies ω_c. The model iteratively chooses its components by creating a set of parameters $P^q = (\gamma_0, \gamma_c, \omega_c)_{c=1}^{q}$, starting with $P^0 = (\gamma_0)$. The iterative process goes as follows:

1) calculate the time series $E(t)$ of errors $E(t_i)$:
$$E(t_i) = R(t_i) - M^q(t_i),$$
2) find γ_{max}, where $|\gamma_{max}| = max\{|\gamma|\}_{\omega\in\Omega}$ and
$$\gamma = avg\left(\left(E(t_i) - \overline{E(t)}\right)e^{jt_i\omega}\right) \tag{3}$$
3) expand the actual model's parameters
$$P^{q+1} = (P^q, \gamma_{max}, \omega_{max}),$$
4) if the termination criterion (see Section III-C) is not met, repeat.

The proposed iterative method needs a suitable algorithm to search for γ_{max} that includes a termination criterion. Therefore, the original question lying in the background of FreMEn of 'what set of frequencies the user needs to choose' is changed to 'what algorithm to search for the global maximum of a function to use'. In the next section, we propose a relatively straightforward search algorithm.

C. Search Algorithm

Although the search algorithm should look for an angular frequency ω, we design and describe this as a search for a *period* P. The relationship between them is given by $P = \frac{2\pi}{\omega}$. We believe it is more convenient for the reader, e.g., considering an angular frequency $\omega = 0.000010389s^{-1}$ and a corresponding period of one week. For a similar reason, we will also use the term *amplitude* for $|\gamma|$.

The graph of amplitudes over the domain of (uniformly distributed) angular frequencies shows a high number of local extremes (Figure 2) related to the parasitic frequencies. That led us to the idea of an iterative sampling-based search

Algorithm 1: Search algorithm – an iterative approach to find the best frequency element

Input:
values - data values,
times - data timestamps,
SR - success rate,
N - number of samples,
NH - sample neighbourhood size,
Iters - maximum number of iterations,
$P_{suspect}$ - suspicious predictions from the previous initialisation (optional)

1 $P_{max} \leftarrow$ half of the *times* duration;
2 $P_{min} \leftarrow$ twice the median time step;
3 $\mathbf{P} \leftarrow rnd^2(P_{min}, P_{max})$ N–times OR $P_{suspect}$;
4 $N_{succ} \leftarrow 0$;
5 **repeat**
6 $\quad |\gamma| \leftarrow$ amplitudes corresponding to periods in $\mathbf{P}$;
7 $\quad |\bar{\gamma}| \leftarrow \frac{1}{N}\sum_i^N |\gamma|_i$;
8 $\quad N_{succ} \leftarrow 0$;
9 $\quad Idx \leftarrow 0$;
10 $\quad$ **for** $i \leftarrow 0, N$ **do**
11 $\quad\quad scale \leftarrow \lfloor |\gamma|_i / |\bar{\gamma}| \rfloor$;
12 $\quad\quad$ **if** $scale > 1$ **then**
13 $\quad\quad\quad N_{succ} \leftarrow N_{succ} + scale$;
14 $\quad\quad\quad$ **for** $s \leftarrow 0, scale$ **do**
15 $\quad\quad\quad\quad radius \leftarrow P_i / NH$;
16 $\quad\quad\quad\quad P_{Idx} \leftarrow rnd^2(P_i - radius, P_i + radius)$;
17 $\quad\quad\quad\quad Idx \leftarrow Idx + 1$;
18 $\quad$ **for** $i \leftarrow Idx, N$ **do**
19 $\quad\quad P_i \leftarrow rnd^2(P_{min}, P_{max})$;
20 $\quad Iters \leftarrow Iters - 1$;
21 **until** $N_{succ} < SR \cdot N$ *or* $Iters > 0$;
22 $P_{best} \leftarrow$ period related to $\max_{|\gamma|} |\gamma|_i$;
23 $radius \leftarrow P_{best} / NH$;
24 $\mathbf{P} \leftarrow rnd(P_{best} - radius, P_{best} + radius)$ for each N;
25 $|\gamma| \leftarrow$ amplitudes corresponding to periods in $\mathbf{P}$;
26 **return** *frequency description related to* $\max_{|\gamma|} |\gamma|_i$

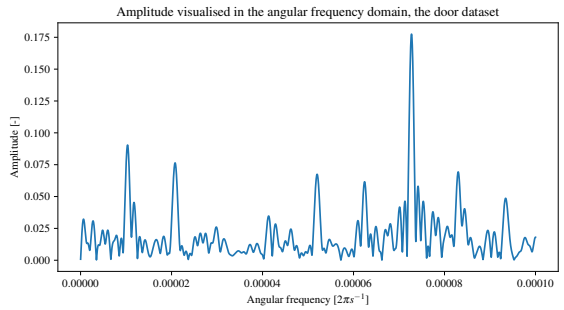

Fig. 2. The graph of amplitudes $|\gamma|$ over the domain of angular frequencies ω using data from the experiment in Section VI-A. The graph shows a high number of local extremes.

for a prominent maximum of amplitudes based on robust consensus, which can avoid local extremes and find one that is dominant in the context of others, see Algorithm 1. From the theory of Fourier transform, the interval of possible periods needs to be bounded from the top by half of the length of the dataset and from the bottom by twice the step [79]. While the calculation of the maximal possible periodicity P_{max} at line 1 is straightforward, estimation of the minimal period P_{min} (line 2) can be tricky when the sampling rate is not consistent. When P_{min} is chosen to be too small, Fourier transform-based algorithms tend to fit the frequency of measurements, which are undoubtedly the most important events in the data but do not provide any information about the measured phenomenon. As the data can also contain a few long gaps in measurements, we estimate P_{min} as twice the median of time differences between subsequent timestamps of the measurements.

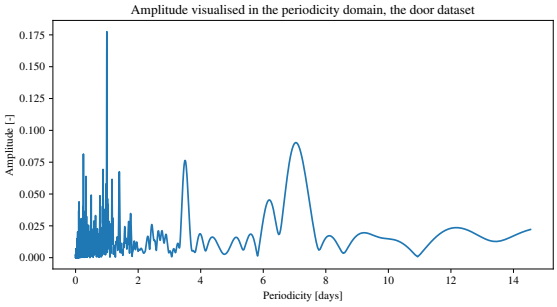

Fig. 3. The graph of amplitudes $|\gamma|$ over the domain of periods P using data from the experiment in Section VI-A. The graph shows many local extremes that form sharp peaks in an area of short periods and wide peaks in an area of long periods. Note that the length of the data equals approximately 70 days.

The graph of amplitudes over the domain of (uniformly distributed) periods (Figure 3) shows that the amplitudes form sharp local maxima at short periods, followed by wider maxima at longer periods. We consider this during the random period selection. The function $rnd(a,b)$ at lines 3, 16, 19, 25 in Algorithm 1 provides a random position *pos* between values a and b such that the random value r between 0 and 1 is generated and $pos = r(b-a)$. Its variant $rnd^2(a,b)$ at lines 3, 16, 19 then produces $pos = r^2(b-a)$ which leads to a preference for shorter periods. At ine 25, we search only for a local maximum, and therefore we apply a uniform random selection $rnd(a,b)$.

After the selection of a set of random periods $\mathbf{P}$, all $P_k \in \mathbf{P}$ are transformed into a set $\{\omega_k = 2\pi/P_k\}$ and the coefficients γ_k of the Fourier series are calculated using equation (3), see line 6. The corresponding amplitudes $|\gamma|_k$ are then compared to their mean $|\bar{\gamma}|$ (line 11), and the subsequent search is partly focused around 'prominent' periods corresponding to 'higher than average' amplitudes (line 15). Note that *scale* at line 11 is an integer, and therefore N_{succ} grows (and the search is targeted) only when $|\gamma|_k$ is more than twice $|\bar{\gamma}|$. N_{succ} cannot grow over $0.5N$, and when it meets its maximum (usually in between $[0.4N, 0.5N)$), it collapses

into a lower value as half of the searches are selected close to a local maximum and the average $|\bar{\gamma}|$ is too high. From the behaviour mentioned above, we can conclude two things:

1) if $N_{succ} \lessapprox 0.5N$, the selection was realised close to a local maximum (line 21)
2) and more importantly, if the algorithm cannot reach a high enough value of N_{succ}, there is probably no prominent periodicity in the data, and, therefore, the search finishes the iterative expansion of its model.

This stopping criterium serves as a natural way to decide on the complexity of the explanatory model. Users do not need to decide on a cross-validation technique to obtain a satisfactory order o, because the generalisation of dynamics of observed phenomenon is an inherent part of the algorithm. Note, that in case a phenomenon is suspected of being random the algorithm returns a constant—an estimation of expected value (Equation 1), or an average of time series values (Equation 2). Similarly, if there is no prominent frequency, i.e., its amplitude is not significantly higher with respect to others, the method also returns an average.

The hyperparameters of the frequency spectrum search, the success rate SR, the number of samples N, the sample neighbourhood size NH, and the maximum number of iterations $Iters$ were tuned by hand on 10% of the 'MHT building lecturer office' of the STRANDS dataset [80]. The data used for hyperparameter tuning were not utilised in the experiments presented in Section IV. The success rate was set to $SR = 0.4$ for the above reasons. The number of samples N directly influences the computational time, which led us to set it to a relatively low value of $N = 100$. We included the supplementary search (lines 22^{+}) for a local maximum to ensure more stable output. The neighbourhood size was set to $NH = 200$, which defines the interval around the period as 5‰ of the length of the period. The maximum number of iterations was set to $Iters = 20$. Reaching the optimal set of hyperparameters was not part of our current research. Regarding the search over the frequency spectrum, the presented approach is relatively straightforward and opens new possibilities for the improvement of selecting the best period.

D. Initialisation

While testing the search algorithm, we found that the output is unstable due to the randomness of the search. Therefore we propose a simple initialisation step that consists of multiple searches for prominent periods at a few different (but large and overlapping) subsets of the training data. Every run of the search algorithm produces a set of suspicious periods $P_{suspect}$. Those periods are then inputted into the main run of the searching algorithm and replace the corresponding number of randomly initiated periods $\mathbf{P}$ (line 3 in the Algorithm 1).

The number of suspicious periods is not limited as it is a set of periods found by multiple runs of the search algorithm. Therefore, there exists a possibility that the number of suspicious periods exceeds the number of samples $|P_{suspect}| > N$ in the main search algorithm. The reduction of the set $P_{suspect}$ would need another heuristic which needs to be designed or optimised according to the knowledge of the data or experience with the tested scenarios. To avoid the increase in complexity and the number of interventions with generally unpredictable impacts, we decided that in such a case, the number of samples N in the main search algorithm is increased according to the number of suspicious periods, $|P_{suspect}| > N \implies N \leftarrow |P_{suspect}|$.

In theory, the increase of N can raise the computational time significantly, as there is no parameter to control the upper limit of the number of suspicious periods. However, in our experiments, we did not experience a situation where the number of suspicious periods exceeded twice the default number of samples N. On the other hand, we experienced one situation when the main search algorithm produced thousands of prominent periods: the binary time series consisted of a few thousand zeros and only two ones, the default number of samples was $N = 5$, and the maximum number of iterations was $Iters = 50$. The calculation carried on for several hours and was manually terminated as normal execution time is in the order of seconds.

IV. Evaluation

The evaluation consists of multiple experiments divided into three sections covering one scientific question each:

1) what difference in prediction quality should we expect when we replace the original method with the proposed method in tasks in which we expect the original method to excel,
2) can the proposed method compete with the original one in robotic experiments that were designed to prove the original method's supremacy,
3) and what is the quality of the predictions of the compared methods when applied to phenomena unrelated to the human week-based calendar?

However, the existing benchmarks are limited by the scope of phenomena considered, namely human routines and environmental changes due to the day-night cycle. Therefore, although we hypothesise that the applicability of the proposed method is much wider, we cannot provide a historically verified set of experiments to support it. We substitute the lack of verified robotic experiments by discussing the relative quality of the proposed method in the context of the questions above. We also weigh the price of good generalisation in the specialised tasks with the new capabilities that the method acquired. The overview of all presented experiments can be seen in Table I.

In Section V, we compare the prediction quality of different Fourier transform-based approaches applied to data describing human presence. We focus on the impact of substituting the original approach with the proposed one on the quality of a model, its predictions, and its learning speed. It was shown in previous works [17], [15], [19], [20] that FreMEn-based predictors using the tuned set of angular frequencies Ω_{168} provide high-quality predictions

Source phenomenon	Section	Time between observations [s]	Time span [weeks]	Environment	Sensor
door	V.A	60	3	office	static RGB-D camera
door	VI.A	60	18	office	static RGB-D camera
human	V.C	330	45	office	static RGB-D camera
human	V.D	4800	45	office	static RGB-D camera
human	V.E	330	45	office	static RGB-D camera
human	V.F	4800	110	office	static RGB-D camera
human	V.G	4800	45	office	static RGB-D camera
image features	VI.B	600	53	open office	SCITOS-G5 w. camera
temperature	VII.A	180	180	space	VZUSLAT-1
proton speed	VII.B	14400	14	space	DSCOVR satellite
proton density	VII.B	28800	14	space	DSCOVR satellite

TABLE I

OVERVIEW OF THE DATASETS USED FOR EXPERIMENTAL EVALUATION.

useful for multiple robotic scenarios. Fourier transforms with Ω_{168} are understood as benchmarks indicating the best available prediction quality. We study the dependence of convergence of the proposed method's prediction quality and the original method's prediction quality on the training dataset's length and amount of data. We also introduce the basic characteristics of the proposed method.

In Section VI, we study the applicability of the proposed method in robotic applications by evaluating the quality of predictions of compared methods in a set of experiments established during the ICRA 2017 Workshop on Reproducible Research in Robotics [81] as a benchmark for spatio-temporal models. In previous works, those experiments were dominated by FreMEn-based methods using Ω_{168} or Ω_{24} because the phenomena in the datasets are derived from day-night changes or the weekly-based routines of people. The fundamental question here is whether the proposed method can compete with FreMEn in its dominant area of applicability.

In Section VII, we evaluate the ability of the compared methods to predict phenomena whose changes are not derived from human routines or the alternation of day and night on the Earth's surface. Here, we study the applicability of the proposed method in more general autonomous robot tasks such as space exploration, which are not bound by terrestrial routines. These experiments play a crucial role in our presentation as we believe the autonomous robot needs the ability to build and maintain a model of its environment without prior knowledge.

A. Criteria

We compare the predictions using three different criteria: *mean squared error* (MSE), *coefficient of determination* (R^2), and *expected encounters* (EE) [19], [20]. Although there exists a debate about the acceptability of MSE in robotic spatio-temporal maps [19], [38], we included it as a conventional criterion that was used in original experiments (Section VI).

The second criterion used in our comparison was R^2, a standard statistical criterion used for assessing goodness-of-fit in the regression tasks. It takes values in the range $(\infty, 1]$ and can be interpreted to represent the ratio of variance explained by the model compared to the uninformed (mean) model. The excellent performance of the mean model and poor discrimination of models is one of the arguments against using MSE-related metrics. In our case, we computed R^2 for individual models over testing data, not training, which allowed us to judge both the degree of overfitting and stability of the data-generating process. The mean model necessarily has $R^2 = 0$ for training data and should stay very close to zero unless there is a change in the process or some numerical artefact.

Following the recent research about criteria suitable for this scientific field [20], we also included EE. As the predictions for the binary variable (estimated probability of the detection [9]) are limited by 0 and 1 (see Equation (1)), the predicted 'zeros' and 'ones' form two classes of impossible and certain events. The events inside such classes are not distinguishable and cannot be ordered by the predicted values. Therefore, we followed the instructions in [20] and every value of EE in our experiments was calculated as the median of 9 independent calculations of EE, each with random reordering of all predictions at the beginning of the calculation.

B. Predictors

The most trivial predictor used in our experiments always predicts an average of the training data. It is denoted as *MEAN*. In some experiments, we included a seasonal windows-based approach that uses predictors derived from the histograms, *HISTday* and *HISTweek*. *HISTday* consists of 24 bins and *HISTweek* consists of 168 bins, where each bin has length 1 hour. Histogram-based predictors predict an average value of a matching bin.

All predictors based on Fourier transform include centring the training data around the mean, calculation of the mean of spectral components γ using available libraries [41], [82], [83], choice of a subset of spectral components, and prediction following Equation (1) or (2) for binary or real-valued time series, respectively. The proposed method is denoted as *NUFTsearch*. If the methods use a set of angular frequencies Ω_{168}, we denote them as *NUFTinfo*, because they are informed a priori what set of periods are expected. Other Fourier transform-based predictors derive the set of

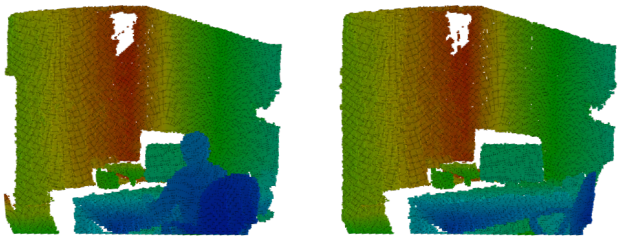

Fig. 4. MHT lecturer office dataset: snapshots of the person present and absent. On the left side of the pictures, we can see an open door (left) and closed door (right).

angular frequencies from the length of data and the median of the steps between consecutive steps. We denote them as *NUFTauto*. The *NUFTauto* method corresponds to the classical Fourier transform, so its complexity, in general, is $\mathcal{O}(n^2)$, but can be sped up to $\mathcal{O}(n\,log(n))$. *NUFTinfo* on the other hand has a fixed amount of frequencies it considers, which makes the computation $\mathcal{O}(n)$. For a similar reason, one can expect the *NUFTsearch* method to have the same complexity as *NUFTinfo* even with the existence of a stopping criterion of the non-deterministic computation, as that is not directly tied to the amount of input data.

All *NUFTinfo* and *NUFTauto* also include a postfix that defines how the order of the model was estimated. *NUFTinfoCV_[criterion]* and *NUFTautoCV_[criterion]* used k-fold cross validation to get the best order o for relevant criterion—*EE* for expected encounters, *R2* for coefficient of determination, and *MSE* for mean squared error. *NUFTinfoM5* and *NUFTautoM5* do not estimate the best order. Their order was set manually to $o = 5$, which makes the methods easy to implement and the model *NUFTinfo* obtained with this order showed good and stable quality of predictions in various human-data-based experiments performed by the authors previously.

V. Examination of Basic Characteristics

FreMEn is usually applied to robotic maps constituted by independent components representing features detected or not at some place and time. Those features form a binary time series; thus, each component has its own model. Those maps then serve as predictors of future states of previously detected features. The whole set of the predicted states of features serves as an estimation of an environment state at the requested time. The robot can then plan its actions according to the expected structure of an environment in the future. If the proposed method is supposed to substitute FreMEn in autonomous robots, we will need to estimate the difference between the original and proposed method's impacts on the robotic maps.

For an initial examination and comparison, we chose the open dataset 'MHT building lecturer office' gathered during project STRANDS [52] that can be downloaded from the project's pages [80]. The data consists of preprocessed video frames, the frame rate is $0.2Hz$, and the length of the video is approximately 2 years. The video was captured in a lecturer's office at the University of Lincoln. Every video frame is represented by a set of values, representing the scene's depth, captured with a 320×240 RGB-D camera, see Figure 4. For privacy reasons, only the 16-bit depth of the image was saved and further processed into a 3D map [72]. A straightforward analysis of part of the depth values determined if the lecturer was available for consultations at the office and if the office doors were open. We transformed the dataset into two binary time series, one tracking the lecturer's presence and one tracking the state of the door (open or closed), and understand them as good representatives of binary robotic map elements. These time series follow a meaningful changing phenomena, the data collection is in line with the usual robotic setup, the sensoric measurements are estimated using a straightforward rule, and the data include real-world noise. Moreover, since the phenomena are directly related to human behaviour (being present in the office and leaving the doors open), we can strongly expect that a model using angular frequencies from Ω_{168} will provide the best forecasts.

We will compare the proposed method with predictors derived from FreMEn using different criteria, training data lengths, and measurement irregularities. Some predictors will use angular frequencies from Ω_{168}, and some will derive the set of angular frequencies from the data. Their order o will be set using k-fold cross-validation over the training data, and, in one case, we will use the default order $o = 5$.

A. Dense and Frequent Data

Our first experiment studies the method's behaviour in the scenario with regular and frequent observation in order to establish and show the performance in the easiest possible setting. For this, we take the data concerning the state of the door every 60 s, following previous works [81], [17]. As noted before, this dataset has been used as a benchmark for spatio-temporal modeling because it is representative of the data the robot collects once it starts performing some task in a long-term autonomy setting. This includes gaps due to the system malfunction and noise on the sensory information.

The original experimental setup and results are provided in Section VI-A. Here, we show the models' forecasts for two selected settings in detail. In these, the models are trained on 28 and 29 days long datasets, respectively, and their forecasts are tested on 7 subsequent days. The lengths of the training datasets were not chosen arbitrarily. We know from previous work that the most distinctive frequencies found by FreMEn refer to one-week and one-day periods, and 28 days correspond to 4 weeks, while 29 days are not divisible by 7. Therefore, the models optimised for week-to-hours-long routines should have an advantage in the 29 days long training scenario against the other ones. The 28 days long scenario, on the other hand, gives uninformed methods the opportunity to provide the best results.

In Figure 5, we can see forecasts of different models. The forecasts are time series of the expected probability of the door being open during the 7 days after the training. The

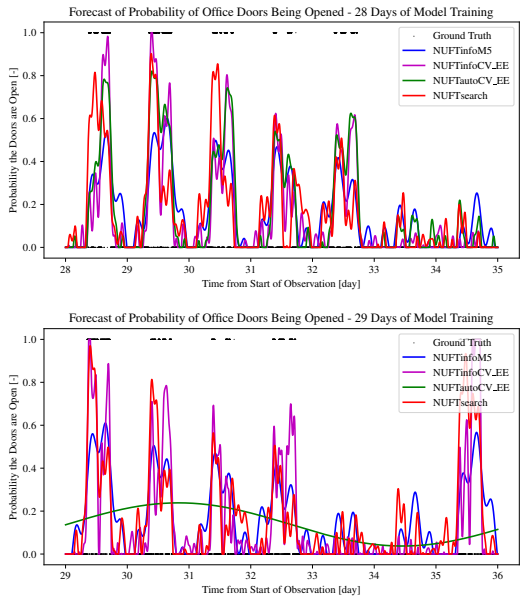

Fig. 5. Forecasts of door state probability for 7 days following the training of the models for 28 days (top) and 29 days (bottom). Both informed models, NUFTinfoM5 (blue) and NUFTinfoCV_EE (purple), and the proposed NUFTsearch (red) provide meaningful forecasts matching the corresponding ground truth (states closed/0 and open/1, black dots) in both scenarios. The forecast of the uninformed model NUFTautoCV_EE in the second scenario is obviously less informative than the other methods due to an unfortunate set of candidate frequencies derived from the length of the training dataset.

measure or quantification of the forecast quality is shown in Table II. For added insight, the table also provides results for MEAN, whose forecasts are not depicted in the graphs. However, in the figure, we can see the ground truth of the experiment, showing when the observed door was open and when it was closed. We can see that the door in the test datasets was open only during the day on weekdays. There is also a visible pattern of the door being closed around noon (lunch). Although the state of doors is, in general, derived from many complex physical, social, legal, and biological rules, such as the orbital characteristics of Earth, job objectives, law on workplace safety, and concentrations of the ghrelin hormone, the presented models are able to generalize them to a certain extent. This generalization and the ability to forecast the probability of the state of such an obstacle can be an advantage [20]—even though the autonomous robot does not actually understand any of the underlying causes.

The 28 days-long scenario results give us insight into the advantage added by the set of candidate angular frequencies Ω_{168}. *NUFTauto* uses a set of candidate frequencies Ω_{auto28}, $\Omega_{168} \subset \Omega_{auto28}$, derived from the length of the training dataset ($\Omega_{auto28} = \{2\pi k/14\,\text{days}, \ldots, 2\pi k/2\,\text{minutes}\}$ and $\Omega_{168} = \{2\pi k/7\,\text{days}, \ldots, 2\pi k/1\,\text{hour}\}$). As we can see in Table II, contrary to $\Omega_{168} \subset \Omega_{auto28}$, the results of cross-validated *NUFTinfoCV_EE* are better then *NUFTautoCV_EE* using any of our three criteria. The results in 29 days-long scenario show apparent failure of *NUFTauto* ($\Omega_{auto29} = \{2\pi k/14.5\,\text{days}, \ldots, 2\pi k/2\,\text{minutes}\}$, $\Omega_{168} \not\subset \Omega_{auto29}$). Its forecast is very close to *MEAN*.

	EE		R^2		MSE	
days of training	28	29	28	29	28	29
NUFTinfoM5	172	159	0.37	0.36	0.077	0.074
NUFTinfoCV_EE	156	135	0.49	0.52	0.062	0.055
NUFTautoCV_EE	162	**517**	0.46	**0.02**	0.066	**0.113**
NUFTsearch	235	187	0.31	0.34	0.084	0.077
MEAN	*711*	*676*	*0.00*	*0.00*	*0.121*	*0.116*

TABLE II

COMPARISON OF RESULTS FOR THE MODELS TRAINED ON THE 28 AND 29 DAYS LONG DATASETS.

NUFTsearch is the worst and second worst of all *NUFT* predictors in the 28 and 29 days long training scenarios, respectively. However, it is not sensitive to the length of the training data, and its forecasts are comparable in quality to the optimized methods. The forecasts of *NUFTsearch* in these two scenarios show that it can be integrated into the autonomous robotic system, and users can expect meaningful results relevant to autonomous decision-making.

B. Moving Observer

Any useful autonomous system needs to first provide a service to its users. Thus, observations the system makes cannot be guided by the need of the technology but rather come as a consequence of its main activity. We assume that the robot is already fully able to perform its task and is equipped with working mapping, localisation, and navigation subsystems.

Such a robot operating autonomously for long periods of time can then observe changes in the state of particular environmental features or obstacles. Although the observed dynamics are not necessarily directly connected to its current task, they can affect the robot routines. If the robot is capable of modelling the dynamics of individual places and building a Temporal MoD [4], it can use this map for better planning or scheduling of its tasks [20] in the future.

To emulate the robot infrequently observing a given location during its deployment, we have subsampled data collected on a continuous basis by a stationary sensor. The chosen environmental feature is the presence of a person in their office, and we assume a robot is operating autonomously, performing an unknown task. The robot moves through the building, deals with its job, and sometimes sees the person in question (observes the environmental feature). The length of the robot's working routine does not correlate with the length of the person's working routine. Its behaviour does not change over time; the person's presence is not an important obstacle that heavily influences the success of its operation. The robot is building a model of the presence of the person that can be useful and included in its map.

The length of the training data gradually grows, and the actual model estimates the person's behaviour in a subsequent week. The forecasts are then compared with the relevant subset of the test dataset obtained from the original

data frames. This experiment shows the convergence of the proposed method and the uninformed methods (methods not using Ω_{168} a priori) with the informed methods, which exploit the a priori knowledge of human behaviour, i.e., which use Ω_{168}. Thus, we will provide fundamental insight into the applicability of the proposed method in applications where FreMEn dominates.

For this experiment, we used the first 217 days of the original data. The data include the summer semester, summer holidays, winter semester and Christmas. We expected that the person's presence in the office during one semester could reflect somehow stable routines of the studied subject, while the summer holidays and the Christmas period will reveal irregularities and changes in the overall dynamics. The behaviour in different semesters probably differs, but the routines are derived from a week-based timetable. Thus, we expect that the process defining the observed subject's presence is, in general, static. One can then see the impact of temporally local violation of this assumption in the graphs attached to the experiments. However, a deeper analysis of these specific periods is out of scope of this study, see basic assumption in Section III-A.

The test dataset consists of detections every 5 seconds and forms a binary time series. The training datasets consist of irregular detections specific to each training data, simulating the robot travelling around the halls and detecting people at specific places.

We use 2 different training datasets consisting of the selection of the measurements from the test dataset. The training datasets simulate a situation where the robot revisits the office on a semi-regular basis with a period consisting of a fixed time representing the length of the patrolling path and variable delay representing unexpected events along the path. Moreover, we model situations where the observation could not be performed, resulting in missing data. We model the variable delay by the exponential distribution and the inability to perform the observation by the uniform distribution. For the sake of simplicity, we did not include the time for charging and other expectable irregularities like maintenance or malfunction of the system. The fixed delay between the measurements in the first training dataset was 5 minutes. The mean value of variable delay was set to 30 seconds, and the probability of missing observation was 10%. In the second experiment, the gap was 1 hour plus an expected delay of 20 minutes and 20% of missing values.

The predictions were evaluated using a hold-out method. Every time the predictors were asked for predictions, they learned the model from the previous data in the training dataset. They predicted the behaviour of the studied phenomenon for a subsequent week. The predictions were consequently compared to the relevant one-week-long subset of the test dataset, and the quality was evaluated using different criteria. For simplicity and visual consistency in graphs, the training time windows were the same in both training datasets. Every time window started with the first measurement and finished at the time of the $i \times 100$-th measurement of the first training dataset, corresponding to a recalculation of the model approximately every 10 hours.

In addition, we adjusted the data for summer time (daylight saving time). Such changes shift the daily human routines by one hour with a frequency of once per year, which is out of the scope of models trained on one year-long data while affecting the predictions. This unnatural, politically enforced change is known in advance, and its detection has no scientific impact.

C. 5 *Minutes with* 30 *Seconds Delay*

The first experiment compares the prediction quality of different models trained by detections gathered approximately every 5.5 minutes; see Figure 6. The top graph shows the comparison using the Expected Encounters criterion [20]. EE was calculated similarly to the original proposal as an ability of a robot to schedule its tasks while avoiding people in their natural environment [19]. In our case, to schedule the task of 'entering the office' while trying to avoid 'meeting the lecturer in his office'. The EE graph provides us with an insight into the lecturer's behaviour. *Theoretical_random* replaces *MEAN*in this graph. It gives us information on the expected number of meetings with the lecturer when the robot does not understand the lecturer's behaviour as a time-dependent phenomenon. The value of the *theoretical_random* corresponds to half of the number of lecturer detections in the one-week-long subset of the test dataset after the prediction time. We can see that the lecturer's behaviour substantially differs throughout the year.

The behaviour of the lecturer mainly follows a week-based timetable as usual in universities. The quality of predictions of the proposed method *NUFTsearch* converges to the original approach *NUFTinfoM5* after roughly 3 weeks. The first semester finishes after 70 days, and we can see that *NUFTsearch* and *NUFTinfoM5* provide similar predictions. Then, for the next 3 months, the lecturer is in his office sporadically, and around day 160, the winter semester starts. The difference between *NUFTsearch* and *NUFTinfoM5* is more considerable in the first half of the winter semester. In the middle graph showing the R^2 criterion, we can see that *MEAN* gives us 'nonzero' values during the first half of the winter semester, which can be interpreted as a change in behaviour. *NUFTsearch* converges to the *NUFTinfoM5* in the EE graph earlier (around day 180) than *MEAN* to value 0 in the R^2 graph (after day 200). The gap in the lecturer's presence after day 250 corresponds to the Christmas period.

The informed variants that use 5-fold cross-validation to estimate the order of the model to achieve the best results in each criterion, *NUFTinfoCV*, expose quite a large room for improvement if the criterion is known in advance. The uninformed variant *NUFTautoCV*, even with 5-fold cross-validation, provides predictions with inconsistent quality. The bottom graph that provides a comparison using traditional MSE cannot be correctly interpreted alone. With the knowledge about the lecturer's behaviour from the EE graph, we can observe that *MEAN* provides the best MSE

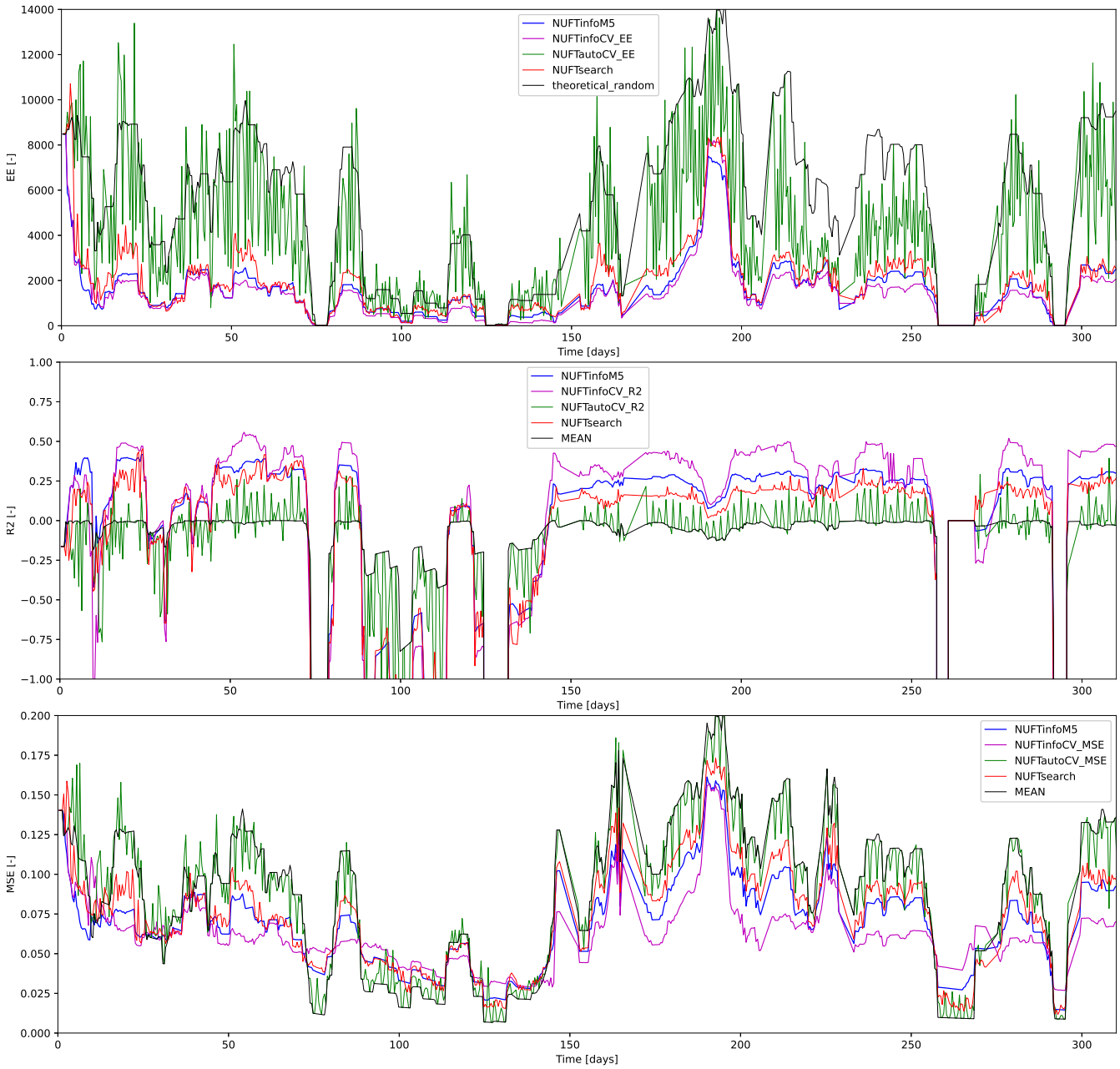

Fig. 6. Human in the office prediction performance in terms of Expected Encounters (EE), coefficient of determination (R^2) and MSE. The *theoretical_random* in EE is drawn from a uniform distribution, thus equivalent to using a MEAN model. Measurements in the training data were taken approximately every 5.5 minute.

in periods during which the lecturer was present less than 5% of the time, which is not a rare situation.

D. 1 *Hour with* 20 *Minutes Delay*

Figure 7 shows predictions of the compared methods trained on detections of the lecturer that were acquired approximately every 80 minutes with every fifth measurement missing. With the length of training data around 14 days, the prediction quality of *NUFTsearch* is very similar to the previous experiment. *NUFTinfoM5* converges to its previous quality more slowly, which gives the proposed method *NUFTsearch* an advantage during the third week. Afterwards, the difference between the quality of predictions when trained on sparse and dense data is not noticeable for both *NUFTsearch* and *NUFTinfoM5*. *NUFTinfoCV* is similar to *NUFTinfoM5*, which probably reflects the lack of fine-grained features of the subject's behaviour in the sparse training data. *NUFTautoCV* provided a higher number of very bad predictions than in the previous experiment.

E. *Identification of periodicities*

In this section, we wanted to compare the periods found by the proposed search algorithm with the original approach; see Figures 8 and 9. The original method uses only a

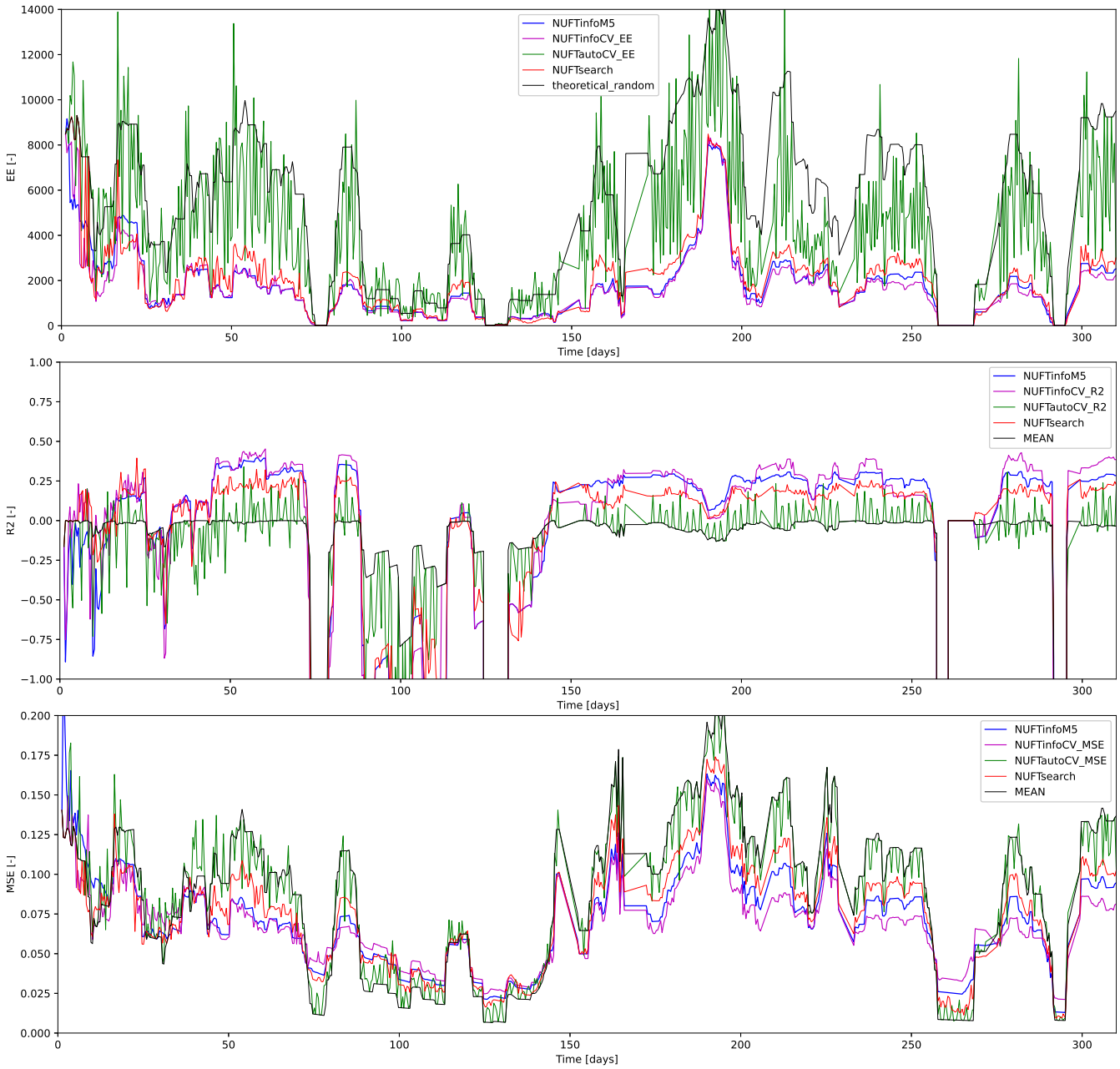

Fig. 7. Human in the office prediction performance in terms of Expected Encounters (EE), coefficient of determination (R^2) and MSE. The *theoretical_random* in EE is drawn from a uniform distribution, thus equivalent to using a MEAN model. Measurements in the training data were taken approximately every 80 minute.

limited but correctly chosen subset of candidate frequencies. It results in fast learning of the most significant periods and the model's stabilisation. On the other hand, the proposed method searches the whole domain for potential periodicities. For example, once the weekly one stabilises, the proposed method found a 'less-than-one-week' periodicity (Figure 9), which cannot be detected by the original approach.

As with most numerical optimisation algorithms, the gain in speed and universality of the new algorithm is traded off for precision. This is best seen in the weekly periodicity, the largest detected and shown in Figure 9. The number of times the algorithm has seen a whole cycle in the data grows relatively slowly. The search algorithm has limited precision in identifying this periodicity and takes a long time to converge.

F. Deterioration of Forecast over Time

The previous section (V-E) showed that the estimation of compositional frequencies by the proposed algorithm does not usually match the human calendar exactly. Such an inaccuracy leads to the deterioration of the forecast over time. To show the impact of this cumulative error, we chose part of the test data depicting the tested subject's stable

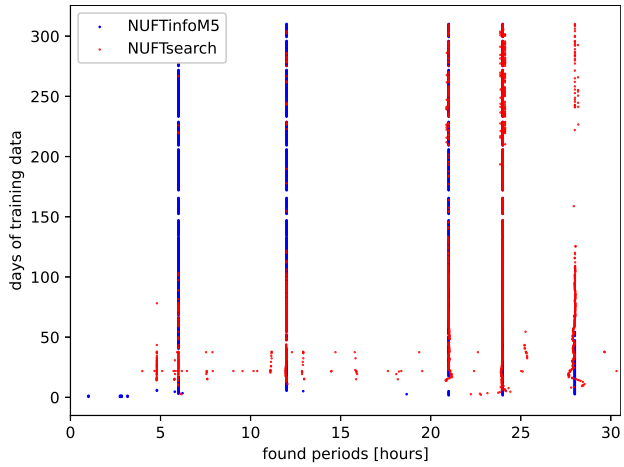

Fig. 8. The performance of the search algorithm with dataset size. On the vertical axis is the time extent of the dataset, and on the horizontal are the periodicities, which in this figure are limited to a little over one day to allow for detail. A point corresponds to an identified periodicity by either our proposed method *NUFTsearch* with no prior information (red) or by the original FreMEn method *NUFTinfo* (blue).

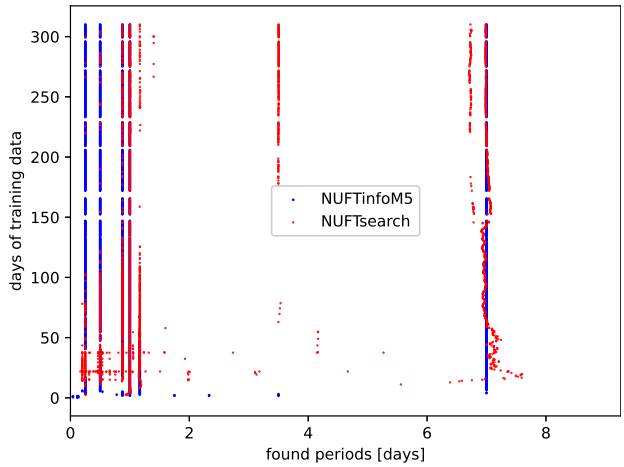

Fig. 9. The performance of the search algorithm with dataset size, similar to Figure 8 but on a larger extent of periodicities. On the vertical axis is the time extent of the dataset, and on the horizontal are the periodicities up to 8 days long. A point corresponds to an identified periodicity by either our proposed method *NUFTsearch* with no prior information (red) or by the original FreMEn method *NUFTinfo* (blue).

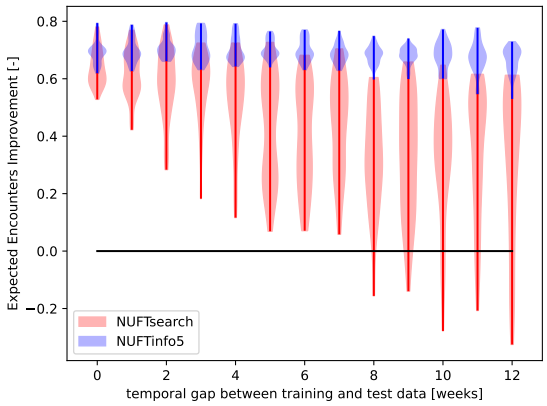

Fig. 10. Deterioration of forecast quality in time. The length of training data is $3Ms$. Event prediction on 1 week long test data is performed after $0-12$ weeks. *NUFTinfoM5* (blue) provides stable predictions for the whole time. *NUFTsearch* (red) provides relatively stable predictions for the first $4-5$ weeks, then the model quality decreases due to cumulative error in periodicities estimation. We detected such models whose prediction quality is worse than *theoretical_random* (black line) after 8 weeks.

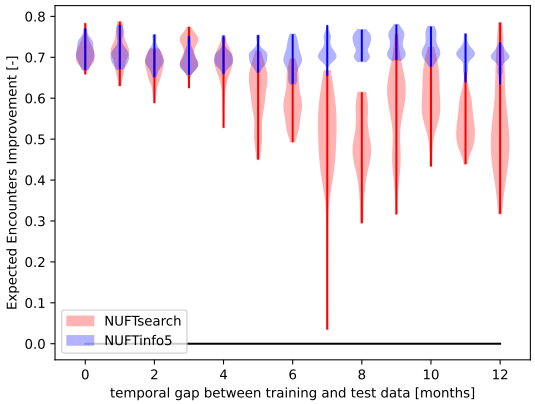

Fig. 11. Deterioration of forecast quality in time. The length of training data is $15Ms$. Event prediction on 1 week long test data is performed after $0-12$ moths. *NUFTinfoM5* (blue) provides stable predictions for the whole time. *NUFTsearch* (red) provides stable predictions for the first $4-5$ months, then the quality of model decreases by approximately 25%. *theoretical_random* is depicted as a black line.

behaviour. The training data are similar to those used in the second experiment (Section V-D). We trained models *NUFTinfoM5* and *NUFTsearch* on constant length parts of the training data with variable temporal gaps between the end of training and the start of the test weeks.

We performed two experiments that differed in the length of training data and the length of the gaps between training and test datasets. The first one with the length of training data $3Ms$ (3 megaseconds, almost 35 days) and temporal gaps $0-12$ weeks, and the second one with the training length $15Ms$ (about 24 weeks) and gaps $0-12$ lunar months (28 days). We trained and tested the models in 80 temporally shifted scenarios in both experiments.

The parameters of the experiments were chosen concerning the easy interpretability of the results. $3Ms$ of training data include 25 working days and $15Ms$ includes 24 working weeks. The estimation of the length of a week after 35 days (5 weeks) of training in Figure 9 looks inaccurate but meaningful. All the training and test parts of datasets in the first experiment (shorter, $3Ms$ training length) lay in the part of data depicting stable behaviour of the tested subject with no gaps in measuring, evading meaningless models and purposeless tests. The number of test scenarios (80) was limited by the length $15Ms$ of training data and the maximal distance between training and test data of almost 1 year.

For the comparison of the quality of forecast over different parts of testing data, we define a novel metric Expected

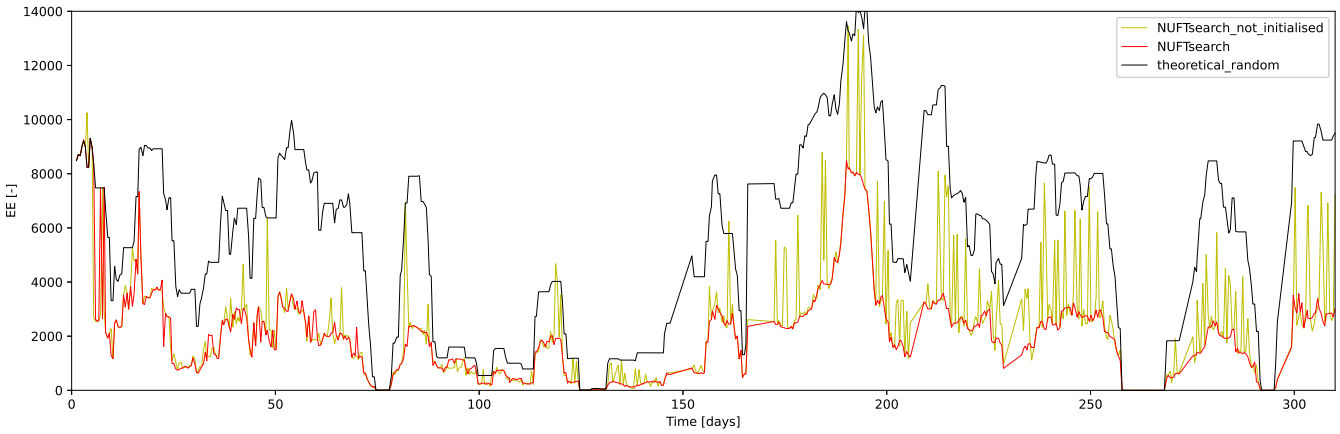

Fig. 12. Difference between using and not using the initialisation. Human in the office prediction performance in terms of Expected Encounters (EE). The *theoretical_random* is drawn from a uniform distribution, thus being equivalent to using a MEAN model.

Encounters Improvement (EEI):

$$EEI_{model} = 1 - \frac{EE_{model}}{EE_{theoretical\_random}}, \tag{4}$$

where EE_{model} is the Expected Encounters of a tested model, and $EE_{theoretical\_random}$ is half of the test dataset's detections, corresponding to a random model. Using such metrics, while assuming, for simplicity, that the 'ideal model' would avoid any encounter, we obtain the improvement of the robot's scheduling of its tasks using the model compared to the scheduling without the knowledge of the environmental dynamics. In other words, the quotient in equation 4 cancels out the number of detections in the test data, allowing us to compare models on different parts of the test dataset.

In Figures 10 and 11, we show the distributions of EEIs calculated for *NUFTsearch* and *NUFTinfoM5* when forecasting $0-12$ weeks and $0-12$ months into the future while training on $3Ms$ and $15Ms$ long training datasets, respectively. The improvement in Expected Encounters of *NUFTinfoM5* is stable in both cases. In the first experiment, Figure 10, *NUFTSsearch*'s improvement of EE is usually lower than EEI of *NUFTinfoM5* but stable for $4-5$ weeks when trained for $3Ms$. We observed some outliers toward deterioration, which means there were a few low-quality models. The deterioration follows our expectations from the previous experiment, Figure 9, where we can see an inaccurate estimation of one-week periodicity when trained on a similarly long training dataset. In the second experiment, Figure 11, EEI of *NUFTSsearch* is similar to *NUFTinfoM5* and stable for $4-5$ months when trained for $15Ms$. After 5 months, the prediction quality decreases by approximately 25%. In our specific case, we observed that the length of forecasting quality's stability matches the training data's length.

G. Note on Initialisation

As seen in Figure 12, adding the initialisation makes the prediction quality of *NUFTsearch* more stable. The frequency of quality drops rises with the training dataset's length. We speculate that the reason lies in the order of the angular frequencies detected as prominent. Less frequent events likely carry a higher error in estimating the correct value of their frequency. When the search algorithm estimates the value of the dominant frequency inaccurately, the estimation of the following frequency is probably influenced by the previous error.

In the previous two experiments, where one can expect weekly routines, we also tried to initialise the main search algorithm with the set of periods equal to the set of angular frequencies Ω_{168} as an alternative to the proposed initialisation, but the quality of predictions was unstable. The reason probably lies in the high number of local maxima at high frequencies, as observed in Figures 3 and 2.

VI. Original Experiments

Here, we compare the performance of the proposed method in the experiments from the original paper [9], which focused on the benefits of real-world robotic deployments. The data covered a real deployment of the robot lasting ten weeks, where the data was gathered continuously by the robot. The evaluated methods, *NUFTinfoM5*, *NUFTsearch*, *NUFTautoM5*, *HISTweek*, *HISTday*, and *MEAN*, vary in their periodicity estimation as described in Section IV-B. The experiments originated from the automatic benchmarking system [81] that was partially updated by experiments found in [17]. The architecture trained many models with different hyper-parameters on the training dataset and selected the best candidate by performing cross-validation on the testing dataset. Then it used pairwise t-tests to establish a ranking of the methods based on their error in prediction. In Figures 13-16, the results of the statistical comparison are visualised in graphs, and when method A shows a statistically significant improvement over method B, there is an arrow pointing from A to B. Note that the graphs are formed by pairwise statistical significance tests using student t-tests. However, due to the issue of alpha inflation when testing multiple pairs, untested assumptions, and the low number of samples, these graphs need to be understood as a complementary, indicative visualisation of results.

A. Door state

In the first experiment, we test the methods' ability to predict a one-dimensional binary variable that represents the condition of a door, whether it is closed or open, located at a university office. An RGB-D camera collected the data describing the door state by checking the occupancy of a specific metric box at the doors. The training dataset was gathered for 10 weeks, and the testing dataset consists of 9 individual subsets, each one week long. The data was collected indoors by a stationary active sensor which utilised structured light to obtain the depth information. Therefore, the data extraction was relatively straightforward compared to a situation where door detection is performed by a robotic platform [84]. The only significant noise in the data was caused by people passing through the opened door, causing false detections of the door state as closed. While the door is open or closed randomly, it is more probable that the door is open during the office hours rather than at night. The probability of the door being open at a particular time was estimated and forecasted by FreMEn [73].

In order to evaluate the prediction capabilities of the methods, we used as the metric the mean square error ε of the predicted state by the individual temporal models $p(t)$ to the testing ground truth state $s(t)$ over all evaluation times in the set T as

$$\varepsilon = \frac{\sum_{t \in T} (p(t) - s(t))^2}{|T|}$$

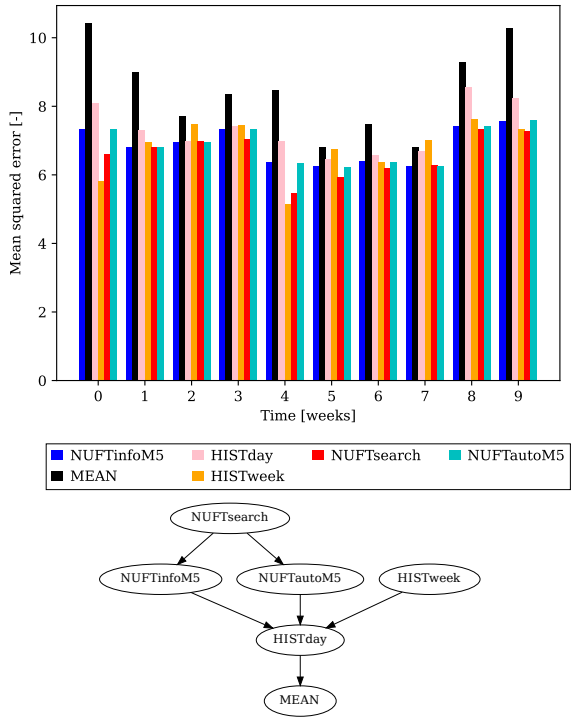

Fig. 13. Door state prediction error. The top figure shows the MSE for the training (week 0) and testing (weeks 1-9) datasets. An arrow from model A to model B in the bottom figure indicates that A's prediction error is statistically significantly lower than for model B.

In Figure 13, the results demonstrate that *NUFTsearch* achieves similar or better results than the state-of-the-art method *NUFTinfoM5*. However, the *HISTweek* method with a bin for every hour in a week also performs similarly to *NUFTinfoM5*. The results of *NUFTsearch* show that even without knowing in advance the present periodicities, the method can successfully learn them and model the temporal behaviour. The high quality of *NUFTautoM5* predictions is based on the fact that the length of the training data is exactly 10 weeks. The set of angular frequencies Ω_{auto} derived from the length of data is a superset of $\Omega_{168} \subset \Omega_{auto}$. Contrary to Section V, the setting of this experiment does not provide correct insight into the difference between *NUFTinfo* and *NUFTauto*. However, we can conclude that it is possible to exchange the original approach with the proposed one without any loss of quality of predictions. Similarly to the conclusion in Section V-F, the deterioration of *NUFTsearch* model trained on 10 weeks long data is not noticeable during the following 9 weeks.

B. Topological localisation

This experiment examines the ability of a robot to localise using only images captured by its onboard camera, given that the robot is also provided with appearance models learnt in advance on images of these locations. The appearance models are helpful, especially if they can capture the variations in the environment [85], [9], [48], [3] because, over time, the appearance of the locations changes as they are located in an open-plan office at a university. This testing scenario aims to test the localisation robustness using the temporal models predicting the occurrence of individual image features in the environment. A mobile robotic platform obtained the visibility information about each feature, a robot SCITOS-G5, which captured images of eight different office locations for one week with a 10 minutes interval between individual images from the same place. Thus the training dataset consists of more than 8000 images. Three testing datasets were collected with the same procedure described but only during one-day-long periods, which yielded 1152 new images for each dataset. Each testing dataset was collected a week, three months and a year after the initial training dataset collection.

The BRIEF descriptor method [86] was used to extract the image features because it achieves high robustness to appearance changes [87]. The temporal models were trained on the features of the same place. The features were matched, and their occurrence in time was assessed from their timestamp. Through this procedure, dynamic appearance-based models were generated for each place; thus, they could estimate the likelihoods of a feature appearing at a specific time.

In order to test the models, the robot uses them to estimate the probability of all features occurring at individual places given the timestamp of the testing image. The robot's current location is estimated by first taking the n features with the highest probability for each place in the office and matching them to the features in the testing image. Then, the place

with the most matches is considered the current location. The applied metric for comparing the tested methods is the percentage of false localisations in the whole testing dataset. In Figures 14-16, there is a relation between the mean error of localisation and the number of features n considered by the methods for localisation.

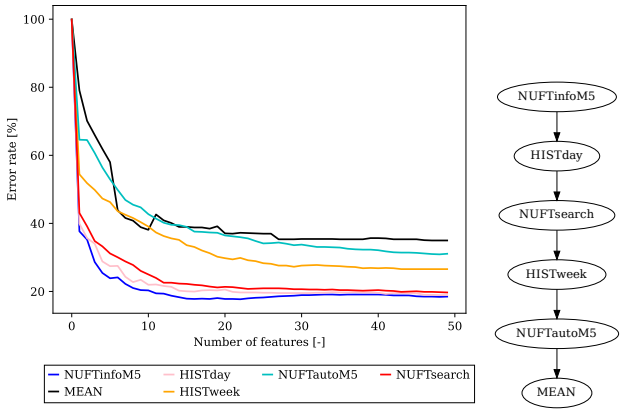

Fig. 14. One day after the training of the models. Temporal model performance for feature-based topological localisation. The left figure shows the dependence of the localisation error rate on the number of features predicted by a given temporal model. An arrow from A to B on the right indicates that A's localisation error rate is statistically significantly lower than that of model B.

The results in Figure 14 present the performance of the methods evaluated on a testing day immediately following the training week dataset. The method *NUFTinfoM5* achieves the predictive capability outperforming all the other methods. *NUFTsearch* together with *HISTday* performed a little bit worse but comparably to *NUFTinfoM5* and distinguishably better than the other methods.

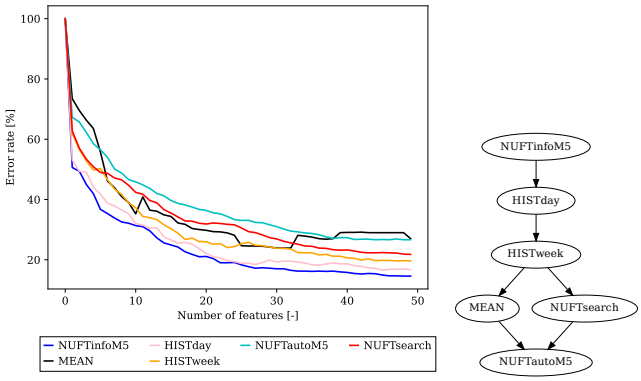

Fig. 15. Quarter of a year after the training of the models. Temporal model performance for feature-based topological localisation. The left figure shows the dependence of the localisation error rate on the number of features predicted by a given temporal model. An arrow from A to B on the right indicates that A's localisation error rate is statistically significantly lower than that of model B.

The testing dataset, evaluated in Figure 15, was collected 77 days (approx. 3.5 months) after the training week. *NUFTinfoM5* and *HISTday* performed better than the others but needed more features to lower their error rate to similar values as in the first scenario. *NUFTsearch* performed significantly worse, and, contrary to the previous scenario, even worse than *HISTweek*.

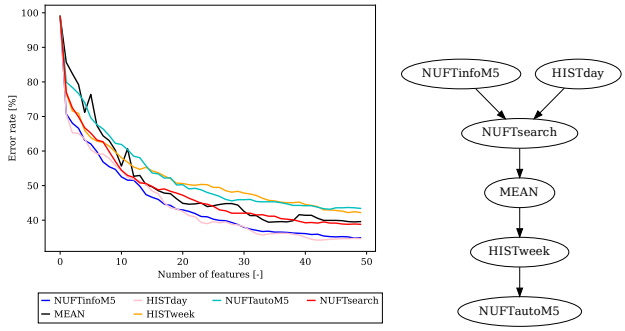

Fig. 16. One year after the training of the models. Temporal model performance for feature-based topological localisation. The left figure shows the dependence of the localisation error rate on the number of features predicted by a given temporal model. An arrow from A to B on the right indicates that A's localisation error rate is statistically significantly lower than that of model B.

The last evaluation was assessed on data gathered for 392 days (approx. a year and a month) after the training period. From the seasonal point of view, we could expect similar features as in the training data. Figure 16 shows that performances of *NUFTinfoM5* and *HISTday* were not discernible and lower than in previous scenarios. *NUFTsearch* was similar to *MEAN* but better than *HISTweek* and *NUFTautoM5*.

We can see that immediately after the training of the models, *NUFTinfoM5*, *NUFTsearch*, and *HISTday* provide similarly good predictions. Because the training dataset is relatively short (only 7 days), the error caused by estimating the frequencies by *NUFTsearch* accumulates, which lowers the quality of the predictions after a more extended period of time. The success of *HISTday* in all three testing scenarios shows that the one-day periodicity is so dominant that it is probably not necessary to model anything else to obtain the high prediction quality in this experiment. However, the change of order in the middle testing scenario (localisation task after 3.5 months) compared to the other two gives an impression that there is another dominant periodicity that needs to be modelled for correct long-term predictions. The hypothetical unknown periodicity is not part of the training dataset and was probably not expected nor targeted by the authors of the experiment. It would be interesting to have much longer data of this kind.

C. Note on Quality of Methods' Comparison

The statistical evaluation of the results raised a few questions. To compare the performance of individual methods, the environment first trained models with different parameters on training data, cross-validated them on testing data, and then chose the best-performing setting for the comparison on the testing data. Such an approach must be considered when implementing methods for comparison, which complicates adding new methods into the framework

for experimental evaluation. As a result, it does not produce a complete picture of the methods' performance, as the method developed for a broader scope will tend to underperform.

It is necessary to point out that the methodology behind the bubble graphs should be reconsidered. First to note is that each experiment uses t-test components and does not consider testing whether any assumptions have been met; much better would be to use the Wilcoxon signed-rank test, which is a non-parametric variant—and, indeed, in subsequent works of the original authors, this was used [88], [89], [90]. The reader might also notice that each experiment is originally performed at a different significance threshold α. The significance levels chosen for the original experiments were 5 and 10 for the "door state" and "topological localisation" experiments, respectively.

This further manifests itself in the last problem such an evaluation has, which is the issue of multiple comparisons [91]. As the test is performed pairwise (and bi-directionally) between all methods (for m methods $n = m^2 - m$ comparisons), the probability of a Type I error in the case the methods are actually performing similarly is $1-(1-\alpha)^n$ for independent tests. In such a case of $m = 6$ similarly performing methods, independent testing on $\alpha = 0.1$ would be expected to get approximately a 95.8% chance of an error and 78.5% chance on $\alpha = 0.05$. To avoid incorrect interpretations of the statistical tests, the bubble graphs are accompanied with charts showing the actual performance of the methods in each case. This can indicate cases where the difference in performance of two particular methods is close to the confidence level.

The statistical tests performed here to create one figure are not independent, but quantifying their dependence would be challenging and is outside this work's scope. In general, such approaches to comparison should be performed with a carefully adjusted α [92].

VII. Analysing datasets with no expert knowledge

This section qualitatively analyses the proposed method and its capabilities in scenarios without prior knowledge of the expected periodicities. All the previous works tested the methods exclusively on robotic datasets from the long-term autonomy domain. It makes sense as that is the domain for which these methods were developed. However, it brings inherent bias because most such applications are usually tested in environments with two significant sources of change, i.e. people and nature on Earth. People and natural cycles then cause the most dominant periodicities to be day-night changes, weekly cycles or other closely related ones. As a result, the methods presented in previous research are typically tailored to detecting these periods.

To eliminate the influence of these two well-known and studied processes, we searched for a domain to test the generality of our methods. We therefore chose the space domain as robotic interplanetary exploration is currently being studied. The space technology domain presented us with an enormous number of possible datasets for testing and getting a better understanding of our methods. By principle, these datasets are not affected by human habits, and the rotation of Earth—which, of course, causes the day-night cycles—can also have no or negligible influence. For these reasons, we searched and chose datasets of phenomena that are naturally periodic, bounded and measured in space. We have no expert knowledge of the context of the data we found. For the sake of simplicity, we test only the *MEAN*, *NUFTinfoM5*, *NUFTautoM5*, and *NUFTsearch* methods in this section.

A. Predicting temperature on board VZLUSAT-1 nanosatellite

First, we experimented with data from the VZLUSAT-1 nanosatellite [93] launched in 2017 to a 510 km low-Earth orbit. The main goal of this mission is as an in-orbit demonstration and testing platform for several new technologies, like specific radiation shielding housing. The satellite is equipped with two sensory payloads, the FIPEX, which is part of the QB50 mission and measures molecular and atomic oxygen density, and Timepix measuring the space radiation along its orbit.

We chose to use the measurements from the temperature sensor of the Timepix experiment [94] as the data are very simple and can be expected to be highly periodic due to the satellites' regular orbit. The dataset [95] contains about three and half years of data, but it is very sparse—it mainly contains day-long bursts of continuous data and then several days with no measurement. We do not know the reason for the missing data, but it provided us with natural splits of the dataset for evaluation. We split every series into the first 70% of data for training and the rest for evaluation; we kept only those where the number of training points exceeded a hundred. The evaluation was done regarding prediction quality measured by standard MSE and R2 metrics.

The results of this experiment are captured in Figures 18 and 19 It is clear that when predicting the data, *NUFTsearch* almost always gives the most accurate predictions with R^2 always greater than 0 but mostly reasonably high. In contrast, the other methods—except for *NUFTautoM5*—do not even get above 0, while simple *MEAN* beats the *NUFTinfoM5* method significantly. It is also worth noting that the situations where the new method does perform generally worse correspond to those where even *MEAN* has poor, negative R^2, which indicates that the data are significantly different to the training situation. To illustrate the behaviour of the methods, we also include an example of training reconstruction and subsequent prediction in Figure 17.

In most cases, the *NUFTsearch* with no prior knowledge identified only periodicities within a minute of 94.5 minutes, which is the period of orbit of the satellite itself. Both methods could detect this because the satellite oscillates between the light and dark sides of the Earth, corresponding to higher and lower temperatures. However, the surprisingly bad results of *NUFTinfoM5* in metrics evaluation are caused

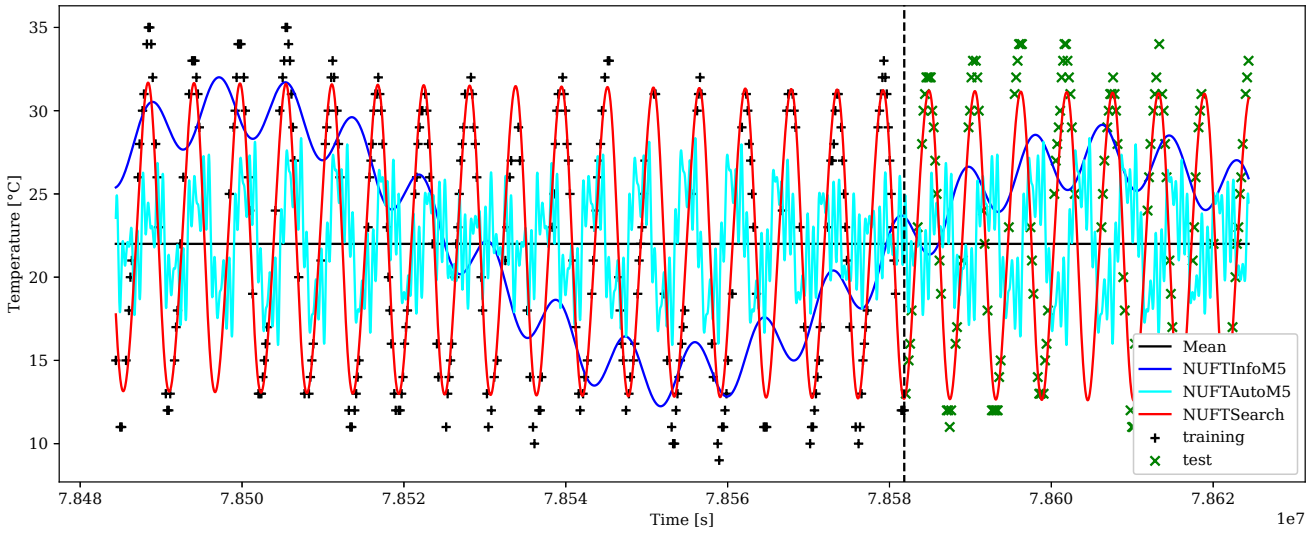

Fig. 17. Example fit and evaluation behaviour of the trained methods. Black crosses represent training samples, the green ones the testing samples, and the dashed vertical line demarks the spit of the time series into training and testing data.

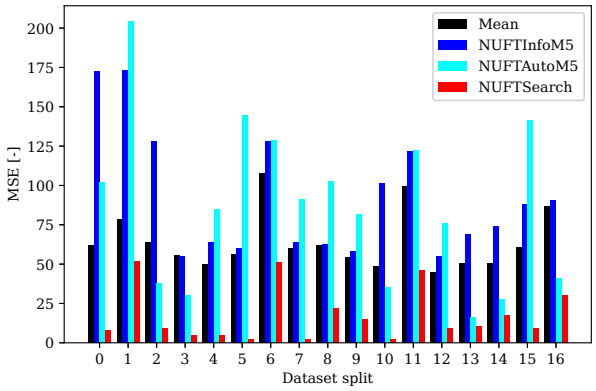

Fig. 18. Results of comparison of MSE in predicting the temperature of Timepix experiment on different data splits.

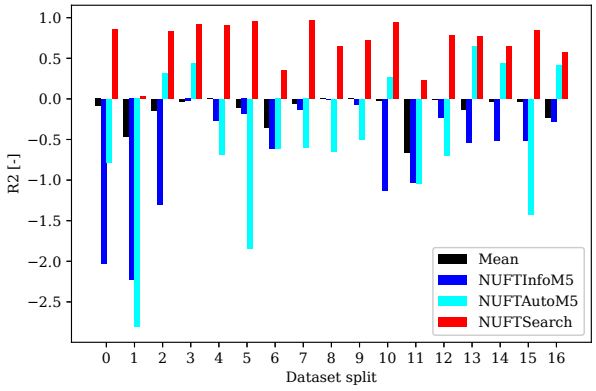

Fig. 19. Results of comparison of R^2 in the prediction of the Timepix experiment's temperature on different data splits.

by it behaving completely chaotically concerning the data, as is also visible in Figure 17.

B. Interpretability of solar wind properties in L1 Lagrange point

The second dataset we tested our methods on was from NOAA's Deep Space Climate Observatory (DSCOVR) satellite [96], which is orbiting the Sun in the L1 Lagrange point—1.5 million km from Earth, located in gravitational equilibrium between Sun and Earth—and provides real-time measurements of solar winds. Its primary purpose is to enable space weather forecasting and provide an early warning system for cases of interplanetary coronal mass ejections, i.e. geomagnetic storms that can be detrimental to the electrical equipment orbiting the Earth as well as to those on the ground. The satellite became operational in 2016 and measured many quantities related to solar winds, like the vectors of the velocity of protons and alpha particles in the wind, the density of these particles and more.

We chose to only deal with protons and try to estimate periods in their speeds and densities. Both datasets represent very different quantities with entirely different scales. The literature on this topic talks about long periodicities we are not used to seeing in human-populated environments. Modelling these phenomena represents the situation where an autonomous system measures and analyses a new, previously unknown quantity of arbitrary periodic behaviour.

The data is sampled every 20 seconds. Analysing such dense data over long datasets demands excessive computational power. As we expect relatively long periods in the data, we aggregated the data by averaging over 4-hour and 8-hour windows for proton speeds and densities, respectively. The window lengths were chosen arbitrarily. The original dataset also contains several gaps. We split it and took the first part, containing approximately one year of data.

	Detected periodicities
NUFTsearch	**13.76** days, 25.75 days, **9.02** days, 13.60 days, **27.36** days
NUFTautoM5	29.88 days, 6.64 days, 14.94 days, 8.54 days, 59.75 days
NUFTinfoM5	**7.00** days, 0.16 days, 0.08 days, 0.17 days, 2.33 days

TABLE III

PERIODICITIES DETECTED IN THE PROTON SPEED. BOLD VALUES ARE THE CLOSEST TO THE ONES PREVIOUSLY ESTIMATED USING A PERIODOGRAM.

	Detected periodicities
NUFTsearch	**13.97** days, **26.19** days
NUFTautoM5	14.88 days, 6.61 days, 29.75 days, **9.92** days, 19.83 days
NUFTinfoM5	**7.00** days, 0.08 days, 0.11 days, 0.11 days, 0.35 days

TABLE IV

PERIODICITIES DETECTED IN THE PROTON DENSITY. BOLD VALUES ARE THE CLOSEST TO THE ONES PREVIOUSLY ESTIMATED USING A PERIODOGRAM.

The question we wanted to answer using these datasets is whether the tested methods can detect meaningful and interpretable periodicities of unknown phenomena. The interpretability of the model's parameters allows us to predict and learn useful information about some periodic phenomena. As periodicities are notoriously hard to learn using, e.g. gradient descent methods, our method can also be used to give other machine learning algorithms reasonable priors.

The solar wind research on large datasets provided information about 7, 9, 14, and 27 days long short-term periodicities [97], which were estimated using a periodogram. In Tables III and IV, we provide lists of detected periodicities in the data by each method. The bold font highlights the results closest to the prior known frequencies. We can see that while *NUFTsearch* did find a variety of periodicities, *NUFTinfoM5* found only one period—one week, which is by coincidence part of its prior. The periodicities detected by *NUFTsearch* correspond to 3 and 2 out of 4 known periodicities in proton speed and density, respectively. *NUFTsearch* did not detect the 7 days long period for some unknown reason. The results show that the model's parameters are explainable.

VIII. DISCUSSION

A. Temporal Structures

In recent years, the main objective of papers that replace stationary probability with a Fourier transform-based time series modelling of events in robotic maps was to demonstrate that using periodicities is beneficial for the autonomous planning of different robotic tasks. Although the benefit was evident in the long-term deployment of robots in human environments [52], the theoretical results in the experiments were, sometimes, not too sound. For example, in [17], the authors compare the prediction quality of FreMEn versus the average values over the spatio-temporal map of human flows with success. However, the error reduction was between 0 and 0.1%. Moreover, the model that predicts 0 everywhere was omitted. The reason was studied in [38] and discussed in [19]—when a human is understood as a point in space and time, the event, human detected at coordinate (x,y,t), is very rare compared to the large number of cells in a $4D$ spatio-temporal map. The authors of [19] therefore proposed a new type of criterion, Expected Encounters, that compares the predictive quality of methods based on their ability to plan robotic tasks, i.e. to minimise the number of potential collisions with humans in this case. Their results still showed the superiority of FreMEn-based methods, mainly STeF-Map [14].

As of now, we believe that the question 'whether the modelling of periodic events in human-populated areas is beneficial for deployed autonomous robots' is successfully answered. The authors of [20] claimed that the human-populated environment does not contain only spatial structures like walls but also temporal structures derived from human routines. Those routines are based on work management, directly connected to the calendar and clock. Similarly to any worker, it is beneficial for a robot to understand the concept of calendar and clock. In the case of the original implementation of FreMEn, this motivated the usage of the set of angular frequencies $\Omega_{168} = \{2\pi k/604800\}_{k=1}^{168}$ (Section II-C). We speculate that there are different temporal structures, not only the one derived from the weekdays; we believe that autonomous robots should have the ability to estimate a "local calendar and clock"' that is natural for the environment in which they are deployed.

B. Toward Autonomous Robotics Paradigm

Defining our work is the focus on the long-term deployment of autonomous mobile robots in natural environments that are not designed for robots' operation. One of the issues such robots have to face is the presence of the aforementioned temporal structures. The temporal structures can be modelled by Temporal MoDs, which have proved their usability in the last decade. A crucial part of FreMEn is its ability to generalise over cyclostationary processes—observing changes in the environment and building models that can forecast the expected state of the environment in future. The generalisation directly depends on its ability to find a specific set of dominant frequencies in the observed dynamics, which depends on a carefully chosen set of candidate angular frequencies—in other words, on domain knowledge of the designers. Moreover, the number of selected dominant frequencies, order o, which defines the complexity of the explanatory model, also relates to the performance and is affected by the designer's expectations or the method's implementation. These parameters lead to environment-specific implementations of an otherwise autonomous building of a suitable environmental model.

In this work, we focus on the estimation of the set of dominant angular frequencies directly from data without designers' intervention. The search for the dominant frequencies in the relevant robotic data is not straightforward due to irregularities in the measurements and the fact that the length of training data is neither related to the frequency of the observations nor the frequencies in the data. We

propose an iterative sampling-based search for a frequency dominant with respect to other frequencies. The iteration terminates when the search cannot assess a new prominent frequency. The proposed approach avoids local extremes in amplitudes, provides a termination criterion, and has asymptotic complexity $O(n)$. A side effect of the spectral features being derived directly from the data is that the algorithm can be applied to time series with any (but consistent) time units. Such facilitation, together with the automatic nature of the model parameters estimation, then makes it more robust to errors brought by human operator expectations and directly applicable to autonomous robots operating in unaccustomed environments.

The proposed search algorithm has multiple hyperparameters that were set manually. The search algorithm was evaluated only in a limited set of experiments and, therefore, is presented as a proof of concept. Still, it proved effective in our experimental evaluation. Using the method, the robot can learn the temporal structure of its environment only from its own observations, without any external input from a human. Incorporating temporal structures into a robot's system makes the robot more integrated into the environment and more effective in providing its service. We believe that the proposed approach—the search for dominant frequencies—is a step forward in the autonomous robotics paradigm, as a robot can discover temporal patterns with frequencies not expected by the robot designer. Therefore, it would allow robots to be deployed in environments with unknown, but cyclic temporal dynamics.

C. Assumptions and Limitations

The contribution of this paper also consists of a deep analysis of the spectral decomposition used in FreMEn. We described its assumptions, identified its limitations, and proposed an alternative approach that overcomes the identified shortcomings. Here, we highlight the assumptions of our method and the limitations we are aware of as follows:

1) The method can model only cyclostationary processes that do not change during the robot's deployment. This is a strong assumption when considering, for example, biological agents such as pedestrians but also animals [98], [99]. Since such applications were also shown to improve the robot's operation, we know that this assumption can be violated under some conditions, but we have not studied this in detail.
2) The method considers only candidate periodicities between twice the usual time between observations and half the data length, which is related to the Nyquist–Shannon sampling theorem. Therefore, the frequency spectrum of recognisable processes is limited by the frequency of observations and the duration of their collection.
3) Due to the lack of appropriate datasets, we have not tested the proposed method on processes consisting of multiple dominant frequencies with a ratio greater than 2 orders of magnitude.
4) In common with other current prediction methods for enabling long-term autonomy, a general limitation is that our method does not have a deep semantic understanding of the underlying processes causing the observed periodicities, rather it is trying to learn temporal structures from correlations in the data. This might be addressed in future work, e.g. using some kind of explainable AI.

However, a more comprehensive list of the method's limitations will come from its application to more practical problems.

D. Scope for Applications

FreMEn-based methods were applied in environments directly affected by the rotation of the Earth and its illumination from the Sun. In most cases, the second most distinctive generative process in these environments was derived from human working routines which follow the calendar and, as a consequence, strengthen the influence of the Earth's synodic day. The methods based on FreMEn with the set of candidate frequencies derived from the week and day provided overall high-quality forecasts. However, it is not obvious whether such environments also include different dynamics. The reason behind this doubt comes from the two following issues:

1) quite commonly, FreMEn-derived forecasts are successfully compared to MEAN-like (static probability) forecasts—however, in section V-A, the method *NUFTautoCV_EE* (Figure 5, bottom graph) with intentionally unsuitable set of candidate frequencies also provided a better forecast then MEAN, see Table II, and,
2) the usual elements in evaluation are directly bound to humans or illumination (human presence in the office, open or closed office door, visibility of image features and alike).

On the other hand, the proposed search approach allows for modelling dynamics with unexpected spectral features and building a Temporal MoD consisting of dynamics generated by mutually independent processes. We can speculate that using our approach even in university office-like environments can avoid an unexpected error in MoD, which might thwart the long-term deployment of autonomous mobile robots. We can avoid situations caused, e.g. by the industrial plants using different types of rotating shifts that influence human routines, and the original FreMEn has to be reimplemented accordingly to be successful [42]. To conclude, we believe that the application of the proposed approach will be more successful in human-populated environments, for which the original FreMEn was fine-tuned.

We expect relevant applications in places on Earth not directly influence by the Sun's illumination (like deep sea), where anomalous processes dominate (artificial systems, phytotrons or foundries as in [42]), where the dominant processes are unknown or under research (like a colony of social insects), or in extraterrestrial places.

In particular, honeybee hives are currently being studied using robotic systems in the project RoboRoyale [100]. Possible expected periodical phenomena include bee hatching time and oscillating of bee-hive weight derived from changes in the surrounding foliage and the growth of different pests that affect the colonies. Other biological systems might include disease spread, recessive gene manifestation, or human behaviour tied to slow-acting hormones rather than circadian rhythm or working hours.

On the other side of the spectrum of studied phenomena is extraterrestrial exploration, a domain of autonomous robotics that cannot utilise the knowledge of a calendar. For example, prediction of local environmental changes, like solar flares or illumination changes at random space bodies, could save instruments from destructive forces and help to plan tasks of extraterrestrial robots—similar to the robots employed on the planet Earth while taking advantage of the original FreMEn approach.

E. Comments on Results

We propose a generalisation of FreMEn that does not use predefined angular frequencies derived from a human calendar but searches the space of angular frequencies for the most prominent ones in the data. We showed in the experiments that the proposed method needs around a 4 days long dataset of a phenomenon, whose periods are derived primarily from one-day and one-week periods, to provide predictions one week into the future comparable to those of the original method that includes an optimised set of candidates of angular frequencies. The method needed around a 3 weeks long dataset to stabilise and not show negative deviations. The predictions of both methods for one week into the future were comparable on an almost one-year-long dataset. However, in situations of significant changes in the number of detections of the phenomenon, the error of the proposed method compared to the original approach was raised.

The proposed method provided predictions of better quality than the original one in the experiment, in which the training binary dataset was 10 weeks long, the prediction quality was tested on the subsequent 9 weeks, and the phenomenon was also derived from one-day and one-week routines. However, the situation was much more complicated on a week-long binary dataset with the most dominant one-day periodicity. The prediction of the proposed method for the next day was comparable but worse than the original one. Three and thirteen months later, the prediction of the proposed method was obviously worse. It showed the weakness of the dominant periodicity estimation from data. As the periodicity of one day was estimated with some error, the error accumulated, and the quality of the predictions was reduced.

A very different situation was found in the experiments with the data whose phenomenon was not derived from the calendar. The proposed method created an incomparably better model than the original approach and provided satisfactory predictions. The original method was not able to provide better predictions than the mean, which exposed its weakness targeted by the proposed method. Nevertheless, visual analysis of the found periods showed that these periods are interpretable and can provide insight into the phenomenon's behaviour.

The experiments showed that the prediction of the proposed method is usually worse but comparable to the original method when applied to phenomena derived from one-day and one-week periods. The predictions of the proposed method on the different types of datasets were consistently good, while the original method failed. We showed that the proposed improvement of FreMEn increased the variety of environments where an autonomous robot can successfully operate without prior knowledge of the underlying processes and solve a wide variety of tasks similar to those tested in the human environments. Regarding the computational efficiency, we compared our search-based method to the original FreMEn using Python 3.7 and the fiNUFFT library [82], [83]. Although both methods have identical asymptotic complexity $O(n)$, the search-based method was slower in our experiments than the original FreMEn [9] by an order of magnitude. However, it is still computationally tractable—the proposed method processed a time series of $100,000$ values in about 1 second.

IX. Conclusion

Frequency Map Enhancement, its applications, and its variants changed the understanding of robotic maps from static ones that try to suppress the changes in an environment to dynamic ones that provide information about periodic events in an environment and give autonomous robots the ability to exploit such dynamics. We proposed an improvement of this approach that focuses on the automatic search for dominant frequencies in the robot's observations defining the dynamics of the environment. The improvement does not need prior knowledge to uncover the generative processes that form an environment's dynamics, allowing autonomous robots to create spatio-temporal maps of unaccustomed and novel environments.

We evaluated the models provided by the proposed method in three different experiments. The first experiment compared the learning process of the original and proposed version of modelling the dynamics of an environment with known generative processes, giving us a view of the price of generalisation in specialised tasks. The second experiment evaluated the quality of the model given by the proposed method in tasks where the original approach previously proved its dominance. The third experiment consisted of applying the methods to the data whose dynamics were generated by processes unrelated to human working habits or day-night changes, showing the proposed method's ability to model unfamiliar dynamics. We showed that the proposed concept of the automatic search for dominant periods in autonomous robot observations provides comparable models in specialised tasks, while substantially extending the possible applicability of Frequency Map Enhancement.

References

[1] P. Biber and T. Duckett, "Experimental analysis of sample-based maps for long-term slam," *The International Journal of Robotics Research*, vol. 28, no. 1, pp. 20–33, 2009.

[2] G. D. Tipaldi, D. Meyer-Delius, and W. Burgard, "Lifelong localization in changing environments," *The International Journal of Robotics Research*, vol. 32, no. 14, pp. 1662–1678, 2013.

[3] W. Churchill and P. Newman, "Experience-based navigation for long-term localisation," *The International Journal of Robotics Research*, vol. 32, no. 14, pp. 1645–1661, 2013.

[4] T. P. Kucner, M. Magnusson, S. Mghames, L. Palmieri, F. Verdoja, C. S. Swaminathan, T. Krajník, E. Schaffernicht, N. Bellotto, M. Hanheide *et al.*, "Survey of maps of dynamics for mobile robots," *The International Journal of Robotics Research*, vol. 42, no. 11, pp. 977–1006, 2023.

[5] J. Shi and T. P. Kucner, "Learning temporal maps of dynamics for mobile robots," *Robotics and Autonomous Systems*, vol. 184, p. 104853, 2025.

[6] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Transactions on robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.

[7] L. Kunze, N. Hawes, T. Duckett, M. Hanheide, and T. Krajník, "Artificial intelligence for long-term robot autonomy: A survey," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4023–4030, 2018.

[8] D. Kothandaraman, T. Guan, X. Wang, S. Hu, M. Lin, and D. Manocha, "Far: Fourier aerial video recognition," in *European Conference on Computer Vision*. Springer, 2022, pp. 657–676.

[9] T. Krajník, J. P. Fentanes, J. M. Santos, and T. Duckett, "Fremen: Frequency map enhancement for long-term mobile robot autonomy in changing environments," *IEEE Transactions on Robotics*, vol. 33, no. 4, pp. 964–977, 2017.

[10] D. M. Bland, T. I. Laakso, and A. Tarczynski, "Analysis of algorithms for nonuniform-time discrete fourier transform," in *1996 IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 2. IEEE, 1996, pp. 453–456.

[11] D. Wei and J. Yang, "Non-uniform sparse fourier transform and its applications," *IEEE Transactions on Signal Processing*, vol. 70, pp. 4468–4482, 2022.

[12] T. Krajník, J. P. Fentanes, O. M. Mozos, T. Duckett, J. Ekekrantz, and M. Hanheide, "Long-term topological localisation for service robots in dynamic environments using spectral maps," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2014, pp. 4537–4542.

[13] T. Krajník, J. Fentanes, G. Cielniak, C. Dondrup, and T. Duckett, "Spectral analysis for long-term robotic mapping," in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 3706–3711.

[14] S. Molina, G. Cielniak, T. Krajník, and T. Duckett, "Modelling and predicting rhythmic flow patterns in dynamic environments," in *Annual Conference Towards Autonomous Robotic Systems*. Springer, 2018, pp. 135–146.

[15] T. Vintr, S. Molina, R. Senanayake, G. Broughton, Z. Yan, J. Ulrich, T. P. Kucner, C. S. Swaminathan, F. Majer, M. Stachová *et al.*, "Time-varying pedestrian flow models for service robots," in *2019 European Conference on Mobile Robots (ECMR)*. IEEE, 2019, pp. 1–7.

[16] T. Krajník, J. P. Fentanes, M. Hanheide, and T. Duckett, "Persistent localization and life-long mapping in changing environments using the frequency map enhancement," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 4558–4563.

[17] T. Krajník, T. Vintr, S. Molina, J. P. Fentanes, G. Cielniak, O. M. Mozos, G. Broughton, and T. Duckett, "Warped hypertime representations for long-term autonomy of mobile robots," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3310–3317, 2019.

[18] J. P. Fentanes, B. Lacerda, T. Krajník, N. Hawes, and M. Hanheide, "Now or later? predicting and maximising success of navigation actions from long-term experience," in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 1112–1117.

[19] T. Vintr, Z. Yan, K. Eyisoy, F. Kubiš, J. Blaha, J. Ulrich, C. S. Swaminathan, S. Molina, T. P. Kucner, M. Magnusson, T. Duckett *et al.*, "Natural criteria for comparison of pedestrian flow forecasting models," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 11 197–11 204.

[20] T. Vintr, J. Blaha, M. Rektoris, J. Ulrich, T. Rouček, G. Broughton, Z. Yan, and T. Krajník, "Toward benchmarking of long-term spatio-temporal maps of pedestrian flows for human-aware navigation," *Frontiers in robotics and AI*, vol. 9, 2022.

[21] Z. Ge, J. Jiang, and M. Coombes, "A congestion-aware path planning method considering crowd spatial-temporal anomalies for long-term autonomy of mobile robots," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023, pp. 7930–7936.

[22] Y. Wang, Y. Fan, J. Wang, and W. Chen, "Long-term navigation for autonomous robots based on spatio-temporal map prediction," *Robotics and Autonomous Systems*, vol. 179, p. 104724, 2024.

[23] Z. Ge, J. Jiang, M. Coombes, and S. Liang, "Enhancing swift and socially-aware navigation with continuous spatial-temporal routing," *International Journal of Social Robotics*, vol. 17, no. 1, pp. 87–98, 2025.

[24] T. Krajník, J. M. Santos, and T. Duckett, "Life-long spatio-temporal exploration of dynamic environments," in *European Conference on Mobile Robots (ECMR)*. IEEE, 2015, pp. 1–8.

[25] J. M. Santos, T. Krajník, J. P. Fentanes, and T. Duckett, "Lifelong information-driven exploration to complete and refine 4-d spatio-temporal maps," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 684–691, 2016.

[26] J. M. Santos, T. Krajník, and T. Duckett, "Spatio-temporal exploration strategies for long-term autonomy of mobile robots," *Robotics and Autonomous Systems*, vol. 88, pp. 116–126, 2017.

[27] S. Molina, G. Cielniak, and T. Duckett, "Go with the flow: Exploration and mapping of pedestrian flow patterns from partial observations," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 9725–9731.

[28] ——, "Robotic exploration for learning human motion patterns," *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 1304–1318, 2021.

[29] M. Kulich, T. Krajník, L. Přeučil, and T. Duckett, "To explore or to exploit? learning humans' behaviour to maximize interactions with them," in *International Workshop on Modelling and Simulation for Autonomous Systems*. Springer International Publishing, 2016, pp. 48–63.

[30] L. Halodová, E. Dvořáková, F. Majer, J. Ulrich, T. Vintr, K. Kusumam, and T. Krajník, "Adaptive image processing methods for outdoor autonomous vehicles," in *International Conference on Modelling and Simulation for Autonomous Systems*. Springer, 2018, pp. 456–476.

[31] T. Krajník, M. Kulich, L. Mudrová, R. Ambrus, and T. Duckett, "Where's waldo at time t? using spatio-temporal models for mobile robot search," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 2140–2146.

[32] T. Vintr, K. Eyisoy, V. Vintrová, Z. Yan, Y. Ruichek, and T. Krajník, "Spatiotemporal models of human activity for robotic patrolling," in *International Conference on Modelling and Simulation for Autonomous Systems*. Springer, 2018, pp. 54–64.

[33] M. Rektoris, "Anomaly detection in periodic stochastic phenomena," B.S. thesis, České vysoké učení technické v Praze. Vypočetní a informační centrum., 2021.

[34] C. Coppola, T. Krajník, T. Duckett, and N. Bellotto, "Learning temporal context for activity recognition," in *22nd European Conference on Artificial Intelligence (ECAI)*, vol. 285. IOS Press, 2016, p. 107.

[35] F. Jovan, J. Wyatt, N. Hawes, and T. Krajník, "A poisson-spectral model for modelling temporal patterns in human data observed by a robot," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 4013–4018.

[36] M. Hanheide, D. Hebesberger, and T. Krajník, "The when, where, and how: An adaptive robotic info-terminal for care home residents," in *ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '17. New York, NY, USA: ACM, 2017, pp. 341–349. [Online]. Available: http://doi.acm.org/10.1145/2909824.3020228

[37] F. Surma, T. P. Kucner, and M. Mansouri, "Multiple robots avoid humans to get the jobs done: An approach to human-aware task allocation," in *2021 European Conference on Mobile Robots (ECMR)*. IEEE, 2021, pp. 1–6.

[38] F. Kubiš, "Application of spatiotemporal modeling used in robotics for demand forecast," B.S. thesis, Czech Technical University in Prague, 2020.

[39] J. Blaha, "Inferring temporal models of people presence from environment structrure," B.S. thesis, Czech Technical University in Prague, 2020.

[40] Laboratory of Chronorobotics. (2019) Kdynakoupit.cz. Accessed: 2022-11-28. [Online]. Available: https://kdynakoupit.cz

[41] T. Krajník, "Fremen [source code]." https://github.com/gestom/fremen, 2017.

[42] V. H. Bennetts, K. Kamarudin, T. Wiedemann, T. P. Kucner, S. L. Somisetty, and A. J. Lilienthal, "Multi-domain airflow modeling and ventilation characterization using mobile robots, stationary sensors and machine learning," *Sensors (Basel, Switzerland)*, vol. 19, no. 5, 2019.

[43] J. Ulrich, M. Stefanec, F. Rekabi-Bana, L. A. Fedotoff, T. Rouček, B. Y. Gündeğer, M. Saadat, J. Blaha, J. Janota, D. N. Hofstadler *et al.*, "Autonomous tracking of honey bee behaviors over long-term periods with cooperating robots," *Science Robotics*, vol. 9, no. 95, p. eadn6848, 2024.

[44] D. Austin, L. Fletcher, and A. Zelinsky, "Mobile robotics in the long term-exploring the fourth dimension," in *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No. 01CH37180)*, vol. 2. IEEE, 2001, pp. 613–618.

[45] G. D. Finlayson and S. D. Hordley, "Color constancy at a pixel," *JOSA A*, vol. 18, no. 2, pp. 253–264, 2001.

[46] J. Mount and M. Milford, "2d visual place recognition for domestic service robots at night," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 4822–4829.

[47] A. Gawel, T. Cieslewski, R. Dubé, M. Bosse, R. Siegwart, and J. Nieto, "Structure-based vision-laser matching," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 182–188.

[48] D. M. Rosen, J. Mason, and J. J. Leonard, "Towards lifelong feature-based mapping in semi-static environments," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 1063–1070.

[49] S. Lowry, G. Wyeth, and M. Milford, "Unsupervised online learning of condition-invariant images for place recognition," in *Proc. Australasian Conf. on Robot. and Automation*. Citeseer, 2014.

[50] N. Sünderhauf, P. Neubert, and P. Protzel, "Predicting the change–a step towards life-long operation in everyday environments," *Robotics Challenges and Vision (RCV2013)*, p. 17, 2014.

[51] F. Dayoub, T. Duckett, G. Cielniak *et al.*, "An adaptive spherical view representation for navigation in changing environments," in *European Conference on Mobile Robots-ECMR 2009*, 2009.

[52] N. Hawes, C. Burbridge, F. Jovan, L. Kunze, B. Lacerda, L. Mudrova, J. Young, J. Wyatt, D. Hebesberger, T. Kortner *et al.*, "The strands project: Long-term autonomy in everyday environments," *IEEE Robotics & Automation Magazine*, vol. 24, no. 3, pp. 146–156, 2017.

[53] T. Krajník, *Long–Term Autonomy of Mobile Robots in Changing Environments*. CTU Press, 2018.

[54] K. Van Laerhoven, D. Kilian, and B. Schiele, "Using rhythm awareness in long-term activity recognition," in *12th IEEE International Symposium on Wearable Computers(ISWC)*. IEEE, 2008, pp. 63–66.

[55] U. Blanke and B. Schiele, "Daily routine recognition through activity spotting," in *International Symposium on Location-and Context-Awareness*, 2009, pp. 192–206.

[56] Z. Zhou, D. S. Matteson, D. B. Woodard, S. G. Henderson, and A. C. Micheas, "A spatio-temporal point process model for ambulance demand," *Journal of the American Statistical Association*, vol. 110, no. 509, pp. 6–15, 2015.

[57] F. L. Bayisa, M. Ådahl, P. Rydén, and O. Cronie, "Large-scale modelling and forecasting of ambulance calls in northern sweden using spatio-temporal log-gaussian cox processes," *Spatial Statistics*, vol. 39, p. 100471, 2020.

[58] R. Senanayake, O. Simon Timothy, and F. Ramos, "Predicting spatio-temporal propagation of seasonal influenza using variational gaussian process regression." in *AAAI*, 2016, pp. 3901–3907.

[59] Z. Zhou and D. S. Matteson, "Predicting melbourne ambulance demand using kernel warping," *The Annals of Applied Statistics*, vol. 10, no. 4, pp. 1977–1996, 2016.

[60] Ç. Ak, Ö. Ergönül, İ. Şencan, M. A. Torunoğlu, and M. Gönen, "Spatiotemporal prediction of infectious diseases using structured gaussian processes with application to crimean–congo hemorrhagic fever," *PLoS neglected tropical diseases*, vol. 12, no. 8, p. e0006737, 2018.

[61] S. H. Garrido Mejía *et al.*, "Predicting crime in bogota using kernel warping," Master's thesis, Uniandes, 2018.

[62] J. S. M. Pabón, M. D. Rubio, Y. Castaño, A. J. Riascos, and P. R. Díaz, "A manifold learning data enrichment methodology for homicide prediction," in *2020 7th International Conference on Behavioural and Social Computing (BESC)*. IEEE, 2020, pp. 1–4.

[63] Z. Zhou and D. S. Matteson, "Predicting ambulance demand: A spatio-temporal kernel approach," in *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, 2015, pp. 2297–2303.

[64] A. Gilardi, R. Borgoni, and J. Mateu, "A spatio-temporal model for events on road networks: an application to ambulance interventions in milan," *Preface XIX 1 Plenary Sessions*, p. 702, 2021.

[65] S. Nilsang and C. Yuangyai, "Activity detection for multi-factors of ambulance demand areas: A case study in bangkok," in *AIP Conference Proceedings*, vol. 2397, no. 1. AIP Publishing LLC, 2021, p. 020001.

[66] V. C. Guizilini and F. T. Ramos, "A nonparametric online model for air quality prediction." in *AAAI*, 2015, pp. 651–657.

[67] F. Ramos and L. Ott, "Hilbert maps: scalable continuous occupancy mapping with stochastic gradient descent," *The International Journal of Robotics Research*, vol. 35, no. 14, pp. 1717–1730, 2016.

[68] R. Senanayake and F. Ramos, "Bayesian hilbert maps for dynamic continuous occupancy mapping," in *Conference on Robot Learning*, 2017, pp. 458–471.

[69] A. Tompkins and F. Ramos, "Fourier feature approximations for periodic kernels in time-series modelling," in *AAAI Conference on Artificial Intelligence*, 2018.

[70] ——, "Periodic kernel approximation by index set fourier series features," in *Uncertainty in Artificial Intelligence*. PMLR, 2020, pp. 486–496.

[71] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, 2013, software available at http://octomap.github.com. [Online]. Available: http://octomap.github.com

[72] T. Krajník, J. Santos, B. Seemann, and T. Duckett, "Froctomap: An efficient spatio-temporal environment representation," in *Advances in Autonomous Robotics Systems*, 2014, p. 269.

[73] T. Krajník, J. P. Fentanes, J. Santos, K. Kusumam, and T. Duckett, "FreMEn: Frequency map enhancement for long-term mobile robot autonomy in changing environments," in *ICRA Workshop on Visual Localization in Changing Environments*, 2015.

[74] S. Bagchi and S. K. Mitra, *The Nonuniform Discrete Fourier Transform*. Boston, MA: Springer US, 2001, pp. 325–360. [Online]. Available: https://doi.org/10.1007/978-1-4615-1229-5_7

[75] T. Krajník, T. Vintr, G. Broughton, F. Majer, T. Rouček, J. Ulrich, J. Blaha, V. Pěčonková, and M. Rektoris, "Chronorobotics: Representing the structure of time for service robots," in *Proceedings of the 2020 4th International Symposium on Computer Science and Intelligent Control*, 2020, pp. 1–8.

[76] T. Krajník, "Spatio-temporal representations for long-term mobile robot autonomy," oral presentation, Habilitation lecture at CTU, Prague, 2019.

[77] Z. Lin, J. Yu, L. Zhou, X. Zhang, J. Wang, and Y. Wang, "Point cloud change detection with stereo v-slam: dataset, metrics and baseline," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 12 443–12 450, 2022.

[78] Y. Mitsuki and T. Kanji, "Diagnosing deep slam for domain-shift localization," in *2022 IEEE/SICE International Symposium on System Integration (SII)*. IEEE, 2022, pp. 724–729.

[79] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 4, pp. 623–656, 1948.

[80] T. Krajník, J. P. Fentanes, C. Dondrup, J. Santos, M. Hanheide, and T. Duckett. (2013) Strands public datasets. Accessed: 2022-11-28. [Online]. Available: https://lcas.lincoln.ac.uk/nextcloud/shared/datasets/index.html

[81] T. Krajnık, M. Hanheide, T. Vintr, K. Kusumam, and T. Duckett, "Towards automated benchmarking of robotic experiments," in *ICRA Workshop on Reproducible Research in Robotics*, 2017.

[82] A. H. Barnett, J. Magland, and L. af Klinteberg, "A parallel nonuniform fast fourier transform library based on an "exponential of

semicircle" kernel," *SIAM Journal on Scientific Computing*, vol. 41, no. 5, pp. C479–C504, 2019.

[83] A. H. Barnett, "Aliasing error of the exp(β1- z2) kernel in the nonuniform fast fourier transform," *Applied and Computational Harmonic Analysis*, vol. 51, pp. 1–16, 2021.

[84] S. Jeon, M. Kim, S. Park, and J.-Y. Lee, "Indoor/outdoor transition recognition based on door detection," in *2022 19th International Conference on Ubiquitous Robots (UR)*. IEEE, 2022, pp. 46–49.

[85] P. Neubert, N. Sünderhauf, and P. Protzel, "Superpixel-based appearance change prediction for long-term navigation across seasons," *Robotics and Autonomous Systems*, vol. 69, pp. 15–27, 2015.

[86] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," in *European conference on computer vision*. Springer, 2010, pp. 778–792.

[87] T. Krajník, P. Cristóforis, K. Kusumam, P. Neubert, and T. Duckett, "Image features for visual teach-and-repeat navigation in changing environments," *Robotics and Autonomous Systems*, vol. 88, pp. 127–141, 2017.

[88] G. Broughton, P. Linder, T. Rouček, T. Vintr, and T. Krajník, "Robust image alignment for outdoor teach-and-repeat navigation," in *2021 European Conference on Mobile Robots (ECMR)*. IEEE, 2021, pp. 1–6.

[89] T. Rouček, A. S. Amjadi, Z. Rozsypálek, G. Broughton, J. Blaha, K. Kusumam, and T. Krajník, "Self-supervised robust feature matching pipeline for teach and repeat navigation," *Sensors*, vol. 22, no. 8, p. 2836, 2022.

[90] Z. Rozsypálek, G. Broughton, P. Linder, T. Rouček, J. Blaha, L. Mentzl, K. Kusumam, and T. Krajník, "Contrastive learning for image registration in visual teach and repeat navigation," *Sensors*, vol. 22, no. 8, p. 2975, 2022.

[91] J. Neyman and E. S. Pearson, "On the use and interpretation of certain test criteria for purposes of statistical inference: Part i," *Biometrika*, pp. 175–240, 1928.

[92] M. Rubin, "When to adjust alpha during multiple testing: A consideration of disjunction, conjunction, and individual testing," *Synthese*, vol. 199, no. 3, pp. 10 969–11 000, 2021.

[93] M. Urban, O. Nentvich, V. Stehlikova, T. Baca, V. Daniel, and R. Hudec, "Vzlusat-1: Nanosatellite with miniature lobster eye x-ray telescope and qualification of the radiation shielding composite for space application," *Acta Astronautica*, vol. 140, pp. 96–104, 2017.

[94] T. Baca, M. Jílek, I. Vertat, M. Urban, O. Nentvich, R. Filgas, C. Granja, A. Inneman, and V. Daniel, "Timepix in leo orbit onboard the vzlusat-1 nanosatellite: 1-year of space radiation dosimetry measurements," *Journal of Instrumentation*, vol. 13, no. 11, p. C11010, 2018.

[95] T. Báča, "Vzlusat-1 timepix radiation data." [Online]. Available: https://github.com/vzlusat/vzlusat1-timepix-data

[96] NOAA. Dscovr: Deep space climate observatory. [Online]. Available: https://www.nesdis.noaa.gov/current-satellite-missions/currently-flying/dscovr-deep-space-climate-observatory

[97] B. A. Emery, I. G. Richardson, D. S. Evans, and F. J. Rich, "Solar wind structure sources and periodicities of auroral electron power over three solar cycles," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 71, no. 10-11, pp. 1157–1175, 2009.

[98] J. Blaha, J. Mikula, T. Vintr, J. Janota, J. Ulrich, T. Rouček, F. Rekabi-Bana, L. A. Fedotoff, M. Stefanec, T. Schmickl *et al.*, "Effective searching for the honeybee queen in a living colony," in *2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2024, pp. 3675–3682.

[99] J. Blaha, T. Vintr, J. Mikula, J. Janota, T. Roucek, J. Ulrich, F. Rekabi-Bana, L. A. Fedotoff, M. Stefanec, T. Schmickl *et al.*, "Toward perpetual occlusion-aware observation of comb states in living honeybee colonies," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 5948–5955.

[100] M. Stefanec, D. N. Hofstadler, T. Krajn'ık, A. E. Turgut, H. Alemdar, B. Lennox, E. Şahin, F. Arvin, and T. Schmickl, "A minimally invasive approach towards "ecosystem hacking" with honeybees," *Frontiers in Robotics and AI*, vol. 9, 2022.