

Machine learning-based forecasting of firemen ambulances' turnaround time in hospitals, considering the COVID-19 impact

Selene Cerna^{a,*}, Héber H. Arcolezzi^a, Christophe Guyeux^a, Guillaume Royer-Fey^b, Céline Chevallier^b

^a*Femto-ST Institute, University Bourgogne Franche-Comté, UBFC, CNRS, Belfort, France*

^b*Service Départemental d'Incendie et de Secours du Doubs, Besançon, France*

Abstract

When ambulances' turnaround time (TT) in emergency departments is prolonged, it not only affects the victim severely but also causes unavailability of resources in emergency medical services (EMSs) and, consequently, leaves a locality unprotected. This problem may worsen with abnormal situations, e.g., the current coronavirus disease 2019 (COVID-19) pandemic. Taking this into consideration, this paper presents a first study on the COVID-19 impact on ambulances' TT by analyzing historical data from the Departmental Fire and Rescue Service of the Doubs (SDIS 25), in France, for three hospitals. Because the TTs of SDIS 25 ambulances increased, this paper also calculated and analyzed the number of breakdowns in services, which augmented due to shortage of ambulances that return on service in time. It is, therefore, vital to have a decision-support tool to better reallocate resources by knowing the time EMSs ambulances and personnel will be in use. Thus, this paper proposes a novel two-stage methodology based on machine learning (ML) models to forecast the TT of each ambulance in a given time and hospital. The first stage uses a multivariate model of regularly spaced time series to predict the average TT (AvTT) per hour, which considers temporal variables and external ones (e.g., COVID-19 statistics, weather data). The second stage utilizes a multivariate irregularly

*Corresponding author

Email address: `selene_leya.cerna_nahuis@univ-fcomte.fr` (Selene Cerna)

spaced time series model, which considers temporal variables of each ambulance departure, type of intervention, external variables, and the previously predicted AvTT as inputs. Four state-of-the-art ML models were considered in this paper, namely, Light Gradient Boosted Machine, Multilayer Perceptron, Long Short-Term Memory, and Prophet. As shown in the results, the proposed methodology provided remarkable results for practical purposes. The AvTT accuracies obtained for the three hospitals were 90.16%, 97.02%, and 93.09%. And the TT accuracies were 74.42%, 86.63%, and 76.67%, all with an error margin of ± 10 minutes.

Keywords: Firemen, Turnaround time, Emergency medical services, Forecasting, COVID-19, Service breakdown

1. Introduction

It is well known that the Emergency Department (ED) is one of the most crowded departments in hospitals [1, 2, 3, 4, 5], where people seek immediate attention given the severity of their problem and long lines are generated contributing to patient dissatisfaction. What is more, EDs' overcrowding can have a knock-on effect on other external services, such as ambulances. Theoretically, when an ambulance arrives at the hospital with patient(s), the crew transfers the patient's care to the ED staff, completes the necessary reports, and cleans and restocks the ambulance (e.g., replaces stretcher linens) to return on service. This total time an ambulance spends at the hospital for handing over patient(s) is referred to as *turnaround interval* [6], which we interchangeably refer to as the turnaround time (TT) of ambulances throughout this paper.

However, with EDs' overcrowding, there is a problem referred to as ambulance offload delay (AOD), which occurs when ambulances' patients cannot be transferred for immediate care to hospitals' ED [1, 7]. On the one hand, AOD risks patients' life due to delays to receive adequate treatment and/or diagnosis, for example. Besides, AOD affects the emergency medical system (EMS) as their ambulance and staff will be in use for more time. That is, AOD directly

increases the ambulances' TT, which poses a population at risk if other major
20 incidents occur and they cannot attend to them [8, 2, 4, 9].

Fire departments are a key component of civil security that ensures the
well-being of the population. In some countries such as France, they are also
responsible for emergency medical services and have been facing a continuous
increase in the number of interventions over the years, which represents a need
25 for the acquisition of more resources and the optimal reorganization of them [10].
In this study, we analyzed the Departmental Fire and Rescue Service of the
Doubs (SDIS 25) that currently has 71 centers deployed throughout the Doubs
region. In total there are 573 cities and 440 districts. Its main cities are:
Besançon, its capital, where the nearest hospital is Besançon Regional University
30 Hospital Center (CHRU), Montbeliard in the north, near the North Franche-
Comté Hospital (HNFC), and Pontarlier in the south, near the Hospital Center
Intercommunal de Haute-Comté (CHIH).

The SDIS 25 is equipped with specialized engines for each type of interven-
tion and the firefighters have various skills in operating these engines. If there
35 are not enough resources available for a long period to attend to an intervention,
either because of the lack of firefighters, the lack of engines or both, a service
breakdown occurs, i.e., the inability to assist within the time limits, which puts
the safety of a certain area or population at risk. For instance, over the years,
it has been observed that most of the breakdowns have occurred in July, since
40 more people go on vacation during this time, including firefighters, i.e., there
could be an increase in the number of interventions and with reduced personnel
more breakdowns would be generated [11]. If we add to this the impact of a
natural phenomenon, epidemic or pandemic, centers would be more affected, as
well as their resilience if resources are scarce.

45 On 12th January 2020, the World Health Organization (WHO) an-
nounced [12] a novel Coronavirus [13, 14, 15, 16, 17, 18], which was officially
named as 'COVID-19' (coronavirus disease 2019) [19] on the 11th of February.
Further, with the spread of COVID-19 across the world, it was declared as a
pandemic [20] on 11th March 2020. In general, multiple measures have been

50 reported for dealing with this novel Coronavirus across the world [21, 22, 23, 24].
In France, as stated in [25], strengthened surveillance of COVID-19 cases was
implemented on 10th January 2020. In their paper, the authors carefully de-
scribed the real-time surveillance system for the first three cases, which were
detected on 24th January 2020. For the scope of this paper, the first official
55 cases in the Doubs region have been reported on 18th March 2020 as shown in
daily statistics published by Santé Publique France in [26]. Nevertheless, the
SDIS 25 early started to attend interventions in which patients had symptoms
of the disease by 29th February 2020.

Although numerous measures were taken to secure peoples' health and well-
60 being, the COVID-19 pandemic has proven to be quite a challenge, particularly
to the public health sector [27, 28, 24]. For instance, in several parts of the
world, numerous news reported an increment to the ambulances' waiting time
in hospitals. In the Texas county US, there are reports of ambulances waiting
about 4-12 hours as hospitals are crowded with coronavirus cases [29]. There
65 is also a report in France, in which ambulances had to wait about 3 hours in a
hospital [30]. This increment could lead sometimes to the worst-case scenario,
i.e, patients' death, as every minute count in such a situation. This is a case,
which was reported in Australia, where two patients have died for long waiting
time in ambulances outside hospitals [31].

70 1.1. *Description of the intervention process*

The process of dealing with an intervention by SDIS 25 is briefly described
as follows:

- (a) First, an emergency call is received and the required armament is gathered
to go to the scene.
- 75 (b) Once arrived at the scene, if necessary, the victim is taken to the hospital.
- (c) At the hospital, the firefighters wait to transfer the victim to the hospital
and to do the corresponding administrative process.

(d) Finally, they return to their center and are available again to attend other interventions.

80 Each performed step requires a record of the firefighter’s status, which is done manually by pressing a button. The firefighter crew chief makes a record to indicate that they are going to the hospital, another record to indicate when they arrive at the hospital, and another one to indicate that they are leaving it. However, there is the possibility of human error, where the firefighter may
85 have forgotten to record the status and registered it a long time later, or on the contrary, where the firefighter may have accidentally recorded too soon.

In addition, an intervention can have several ambulances departing at different times from the scene, depending on the number of victims, or either, sometimes in the same hour we can have several simultaneous interventions but
90 of different types. If firefighters spend more time in hospitals, i.e., high TT, there will be fewer resources available at the centers to respond immediately to an incident. Notice that the terms ambulance and firefighter’s ambulance are used interchangeably throughout this paper.

1.2. Purpose and contributions

95 With these elements in mind, the present work proposes a novel methodology based on machine learning (ML) to make predictions for the TT of each ambulance in a given time and hospital, aiming to provide a decision support tool for SDIS 25 and EMSs, in general. This way, EMSs can activate various proactive mediations according to the time their personal and resources will be
100 in use, aiming to mitigate service breakdown and, consequently, being able to save more lives. In other words, for the short-term, such predictions could allow better allocation of the remaining and available resources if known a priori the time each ambulance will spend in hospitals. For the medium- to long-term analysis, such a system can be improved to re-calculate a better hospital option considering the predicted TT. To summarize, this article proposes 3 main
105 contributions described in the following.

- (a) *Analysis of the COVID-19 impact on ambulances' TT and the breakdowns generated in the fire service.* We provide an in-depth analysis of the COVID-19 impact on the average turnaround time (AvTT) per day in each hospital during the first semester of 2020 in the Doubs region. Further, we describe the already existing breakdown problems in the fire service due to long TTs in hospitals. Finally, we show how this problem worsens with the arrival of COVID-19, demonstrating the need for a system to prevent breakdowns in the fire service in the face of a pandemic.
- (b) *Creation of a regularly spaced time series model that represents the average hourly turnaround time at each hospital.* The model is based on the history of TTs reported by SDIS 25 ambulances since 2015 for CHRU and CHIH hospitals, and since 2017 for HNFC, in order to recognize hourly, daily and weekly trends. To this was added internal explanatory variables such as the number of interventions by firefighters (a greater increase in interventions in a certain period of time may indicate a greater number of victims attending the hospital) and the suspected cases of COVID-19 registered by the SDIS 25. External variables such as the number of COVID-19 cases officially reported in the region; keywords most searched in Google to retrieve trends; traffic predictions, since a very congested day can generate accidents and victims to transport; and meteorological data were also considered, since many floods usually occur in the region, which generates material damage and victims. Four state-of-the-art ML models were considered in this task, i.e., a decision-tree based model (LGBM - Light Gradient Boosted Machine), a feedforward type neural network (MLP - Multilayer Perceptron), a recurrent neural network (LSTM - Long Short-Term Memory), and a decomposable time series model (Prophet).
- (c) *Creation of a irregularly spaced time series model to predict the turnaround time for each intervention and its ambulances.* One or more ambulances belong to an intervention, if the intervention has a long duration, the ambulances could go to the hospital at different times of the day. It is

here where the first model described in (b) will provide us with predicted information on the AvTT in a given hour as an explanatory variable. Along with the type of intervention (to recognize the degree of severity of the incident) and other variables mentioned previously we propose to forecast the TT for each ambulance. Because in the previous task (b) the LGBM model consistently and considerably outperformed the other ML models, for this task of predicting each ambulance’s TT (not average TT per hour), we performed experiments with only LGBM.

To the authors’ knowledge, this is the first work that investigates the impact of COVID-19 on ambulances’ TT. With our findings, we noticed that SDIS 25 ambulances’ TT increased during the outbreak, which led to more breakdowns in service due to the shortage of ambulances that would normally return to service in less time. This way, aiming at providing a decision support tool for SDIS 25 and EMSs, in general, this paper also proposes a novel ML-based methodology to forecast the TT of each ambulance in a given time and hospital. More precisely, considering that ambulances can arrive at hospitals anytime, our problem (c) refers to irregularly spaced time series [32], i.e., the spacing of observation times is not constant (e.g., per day). Thus, we first propose to reconstruct a regularly spaced time series (i.e., AvTT per hour), which gives us a reference on the average time that several ambulances could have been waiting, and thus, recognize a daily or weekly seasonality. Since not every hour nor every day SDIS 25 ambulances went to hospitals, a linear interpolation was applied to complete the dataset. This is because SDIS 25 is not the only institution to transport victims to hospitals. Finally, besides the prediction of AvTT per hour/hospital (b), with the second model (c), we can refine the predictions for each ambulance, i.e., an irregularly spaced time series problem.

The present paper is organized as follows. Section 1.3 reviews contributions from related works. Section 2 details the construction of datasets (Subsection 2.1); the data analysis performed to understand the normal and abnormal behaviors in the collected data (Subsection 2.2); the breakdowns generated due

to long TT periods (Subsection 2.3); the proposed methodology composed of two forecasting models and their baselines (Subsection 2.4); finally, we describe the metrics used to evaluate our models (Subsection 2.6). Section 3 presents the results of the predictive models. Section 4 discusses this research and its impact in fire brigade services. Finally, Section 5 shows our conclusions and future work.

1.3. Literature Review

ED overcrowding has been reported as one of the main causes of AOD [4, 1, 3, 5] since ED staff may not prioritize ambulances' patients. AOD may lead to several risks to patients' care and directly increases the ambulances' TT in hospitals, which puts a population at risk if other major accidents occur and ambulances are still in use. In [8], the authors conducted a prospective longitudinal study to analyze the impact of ED overcrowding on firefighters' AOD. The study identified 21,240 cases where their ambulances were out of service due to AOD, which may have a significant effect on their ability to respond to other calls. As noticed in the survey work on AOD from Li et al. [7], the impact of AOD on EMS resource availability has received less attention from the research community.

In [6], the authors coined the term *turnaround interval* for the total time an ambulance spends at the hospital, i.e., transferring patient(s)' care, completing paperwork, and reestablishing the equipment for the next call. The authors conducted a prospective study by analyzing firefighters' ambulances' activity on transferring 122 patients to a hospital. In [33], the authors designed a discrete event simulation model to evaluate the change on AOD by having dedicated ED nurses for ambulances' handover. Although the authors identified that this practice may reduce the ambulances' TT, this would also lead to low staff utilization. In [34], the authors investigated the relationship between ambulances' TT with patients' acuity, destination hospital, and time of the day, using one-year data from 61,094 patients. In [2], the authors have analyzed the impact of ED overcrowding and ambulances' turnaround interval.

Notice that most aforementioned works have been performed to highlight the importance of the topic and to identify the relationship between ED overcrowding and ambulances' TT, for example. To the authors' knowledge, there is no research trying to forecast the TT of ambulances in hospitals, as we present in this paper, but rather for the prediction of patients' waiting time in ED [35, 36, 37]. On the other hand, within the context of COVID-19, multiple works have investigated machine-learning-based solutions to help to fight the outbreak, e.g., forecasting of ED volume [38] and forecasting the number of confirmed COVID-19 cases [39, 40, 41, 42]; we refer the readers to recent survey works on ML forecasting models and COVID-19 in [17, 18].

In research directly related to firefighters and their processes, Pirklbauer and Dieter [43] proposed a data-driven forecasting model for the type of firefighter interventions. The models were built using two-years data, which achieved a preliminary accuracy of about 61%. On the other hand, our group has proposed a multi-forecasting model to the number of firefighter interventions per region in [10]. In that work, it was noticed that while the number of interventions is increasing each year, fire departments' resources are reduced. For this reason, there is a need to optimize the use of firefighters' personnel and resources, which could relieve the pressure on this emergency medical transport system. The complete methodology also allows a privacy-preserving publication/sharing of interventions' location data, where the Extreme Gradient Boosting technique was used to forecast the number of interventions per region each day with both original and privatized data. Also, it is necessary to know the current status of the fire department's resources and services in order to develop long and short-term reorganization strategies. Thus, in [11], a methodology was created for the calculation of service ruptures, which allows identifying the types and quantity of ruptures, as well as the unavailability of resources at a given time. In addition, with breakage data calculated in the last 3 years, predictive models were built for the number of daily breakdowns, considering the number of firefighters and available engines as the main variables.

2. Material and Methods

In this section, we describe the collection process of internal and external variables. Besides, we analyze the recorded ambulances' TT through the years and the impact of COVID-19 on ambulances' TT; and, consequently, the impact of COVID-19 and high TTs on the fire service. Further, we present our proposed methodology for forecasting the TT of each ambulance and AvTT per hour, and the metrics used to evaluate our models.

2.1. Data Collection

The main source of data comes from the Departmental Fire and Rescue Service of the Doubs department, France. It contains a history of ambulances arrivals collected from January 2015 to June 2020 for CHRU and CHIH hospitals, and from February 2017 to June 2020 for HNFC. For each hospital, two types of datasets were created:

- Arrivals dataset (Arr-DS). The samples represent the TT of each ambulance in a hospital at any time of the day. The features are: year, month, day, day of the week, day of the year, and hour. Also, an indicator to recognize if it is the beginning or end of the week, month, and year. In addition, it was added three variables that jointly describe the type of intervention, where each variable represents a set of subcategories. The first variable represents the 5 main activities that firefighters cover, for example: rescue people, fires, among others. For each activity, there are subcategories that describe the level of urgency (second variable), for example, in rescue people, we found 6 subcategories, two of them are "Emergency situation" and "Special Circumstances of the emergency (public road)". The third variable reports the state of the victim, for example, in "Emergency situation", there are 10 subcategories, two of which are "Severe exteriorized or external hemorrhage" and "Respiratory distress". Finally, the TT in minutes (min) recorded was included. In some records, we found TTs less than 5 minutes, which is very rare, and it is possible that they were

caused by a bad manual recording. Therefore, if the TT is less than 5 minutes, we clipped it at 5. Similarly, there were five unusual samples with TT between 4-7 hours, which were clipped at 90 minutes, since most TTs (99.7%) had less than or equal to 90 minutes.

- 260 • Average turnaround time dataset (AvTT-DS). This dataset has been organized by hour, where each hour represents the average TT of all ambulances that arrived at the hospital in that hour. For example, if between 15h and 16h there were 3 ambulances with TTs: 20 min, 15 min, and 30 min, the mean was 21.67 min for 15h. In addition, each sample includes
265 the following features: year, month, day, day of the week, day of the year, hour, an indicator to recognize if it is the beginning or end of the week, month, and year. And, since 90% of all hospital arrivals have had an AvTT between 10 and 60 minutes, we set the average minimum and maximum with these values. The hours in which there were no arrivals and,
270 therefore, there was no recorded TT and AvTT, were completed by linear interpolation since the TT over the years has been constantly increasing.

In order to discover possible influential external variables, we built another dataset (External-DS) by hour from “01/01/2015 00:00:00” to “30/06/2020 23:00:00”, where we incorporated variables from the following sources:

- 275 • Google Trends. We used the Pytrends¹ library to get the hourly trends in a scale from 0 to 100 for the keywords in french: ‘H1N1’, ‘coronavirus’, ‘SARS’, ‘Influenza’, ‘COVID-19’, ‘diarrhee’ (diarrhea), ‘grippe’ (flu), ‘varicelle’ (chickenpox), ‘incendie’ (fire), ‘inondation’ (flood), ‘greve’ (strike), ‘samu’, and ‘suicide’.
- 280 • Bison-Futé [44]. This source gives us the prediction of the traffic level for the Doubs region, as indicators from 1 to 4, where 1 means a regular circulation and 4 means an extremely difficult circulation.

¹<https://github.com/GeneralMills/pytrends>

- Météo-France [45]. It provides us with historical weather information such as: precipitation, temperature, barometric trend, pressure, humidity, dew point, wind direction, wind speed and gust speed. 285
- Data Gouv [26]. It is a platform for the diffusion of public data from the French government, from which we extracted information on the situation of the COVID-19 pandemic from March 2020 to June 2020 to the region of Doubs (Department 25). Data prior to this date were completed with zeros, as there were no statistics reported. 290

Other variables included from the primary source were the total number of interventions recorded in a given hour over the 6 years, and the number of cases attended with a suspect of COVID-19 per hour from February 2020 to June 2020. This last variable was completed with zeros before February 2020 too.

2.2. Data Analysis 295

The dataset at our disposal has 78,777 interventions where firefighters transported victims to one of the three aforementioned hospitals. The frequency on the number of times each hospital received a victim is 55.44% (CHRU), 18.27% (CHIH), and 26.29% (HNFC), respectively. In order to understand how the ambulances' TT are distributed in our dataset, Figure 1 illustrates, for each hospital, a histogram with bins of 1 minute and the cumulative number of TT in hours (y-axis) for each day of the week and hour in the day (x-axis). 300

In the first column of Figure 1, CHRU, CHIH, and HNFC have right-skewed distributions with mean and standard deviation (std) values as 18.43 ± 10.68 , 14.91 ± 8.66 , and 22.61 ± 11.26 , respectively. In the second column of Figure 1, one can notice a similar pattern for the cumulative sum of TT in hours that firefighters' ambulances spend in the three hospitals with different peak values, which depends on the frequency of times each hospital received victims. This pattern is also noticed in the works [34, 2]. In our case, from 8h in the morning on, the TT starts to increase and remains high up to 19h when it starts to decrease. Also, between 0-6h, the highest cumulative TT is during the weekend, 310

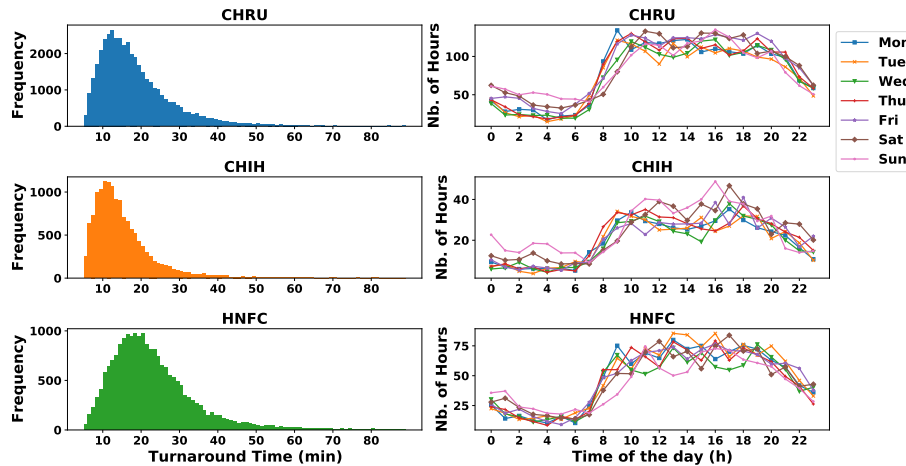


Figure 1: Statistical analysis for the TT of ambulances: each row illustrates a histogram with bins of 1 minute to the TT frequency distribution (left side) and the cumulative sum of TT in hours for each day of the week and hour in the day (right side), considering each hospital. On the left side, the frequencies for the three hospitals follow a positively skewed distribution, where more than 80% of cases were less than one hour. On the right side, we can deduce that between 8h and 22h the accumulated time follows the intense human activity while awake, which is translated as more workload and longer TTs.

i.e., Friday, Saturday, and Sunday. This is because at weekends people tend to go out more and until the wee hours of the morning, which might lead to more accidents, more patient visits to hospitals during on-call hours with reduced staff, and results in a slight increase in ambulances' TT.

In addition, as stated in the introduction (Subsection 1.2), this work aims to study the COVID-19 impact on ambulances' TT in the three aforementioned hospitals. First, Table 1 exhibits for each semester (Sem.) of the analyzed years (2015-2020) and hospital, the following statistics: the total number of arrivals (Arr.) and the mean \pm std TT values in minutes.

As one can notice in Table 1, CHIH has fewer arrivals and a shorter AvTT per semester whereas CHRU presents more arrivals than the other two hospitals since it is located in the capital of the territory, and a lower AvTT than HNFC. Besides, when looking at the number of arrivals for the first semester, it can be detected that there was a higher workload in the firemen department for the

Table 1: Data analysis, for each hospital, on the number of arrivals (Arr.) and Mean±std TT values in minutes per semester during 2015-2020.

Sem.	Year	CHRU		CHIH		HNFC	
		Arr.	Mean±std TT	Arr.	Mean±std TT	Arr.	Mean±std TT
1st	2015	3,114	17.72±10.24	1,230	14.50±7.84	-	-
	2016	3,125	17.68±10.36	1,176	14.33±7.64	-	-
	2017	3,676	17.55±10.31	1,288	13.99±7.43	2,272	22.97±11.12
	2018	5,233	18.75±11.37	1,548	15.05±9.41	3,824	22.91±12.18
	2019	4,672	18.94±11.66	1,368	15.63±9.80	2,994	22.92±11.90
	2020	3,899	20.67±12.57	1,184	15.96±10.51	2,136	24.42±13.19
2nd	2015	2,893	17.59±10.28	1,174	14.69±8.91	-	-
	2016	3,320	17.77±10.32	1,222	13.88±7.32	-	-
	2017	4,829	18.93±11.76	1,556	15.04±9.41	3,454	22.07±10.90
	2018	5,182	18.56±11.35	1,440	15.50±9.32	3,600	22.04±11.55
	2019	3,730	20.08±13.03	1,208	16.52±11.23	2,430	23.25±12.37

years 2018 and 2019 compared to the other years. In the case of 2020, the reduced workload was due to a lockdown period², which decreased the movement of people and reduced the number of firefighters' interventions, for example, fewer traffic accidents. The second semester of each year normally presents a higher AvTT, which could be due to two holiday periods (Jul-Aug and Dec), where people travel more due to vacations. Although the AvTT normally increased for each hospital throughout the years according to its workload, the AvTT for 2020 is higher than all years and for the 3 hospitals, even with a reduced workload compared to 2018 and 2019. For instance, hospital CHRU presented about 9.13% more AvTT in the first semester of 2020 than in the first semester of 2019, which is 9 times higher comparing the increment for the same period for the years 2019 and 2018 (1%). Similarly, hospital HNFC presented about 6.54% more AvTT in the first semester of 2020 than in the first semester of 2019, while the years 2017-2019 presented similar AvTTs. This proves the impact of the pandemic on the increase in ambulances' TT.

Furthermore, we now try to comprehend how the on-going pandemic has

²https://en.wikipedia.org/wiki/COVID-19_pandemic_in_France#Lockdown_measures

made an impact on the ambulances' AvTT per day on hospitals for 2020 only. For this purpose, the plot on the top of Figure 2 illustrates for each day (x-axis) the AvTT in minutes (y-axis) for each hospital. Additionally, at the bottom of this figure, the plot illustrates the official COVID-19 statistics from two data sources, i.e., Data Gouv and the ones reported by the SDIS 25, to search for evident patterns. More specifically, this plot illustrates the current number of hospitalized individuals (hosp.), the number of patients in reanimation or critical care (rean.), the cumulative number of individuals who returned home (ret. home), and the cumulative number of individuals who died (dead) regarding the Data Gouv source, and the number of cases per day attended by the SDIS 25 with a suspicion of COVID-19 (Susp. COVID-19). For both data sources, there is an indication of the first day with COVID-19 cases.

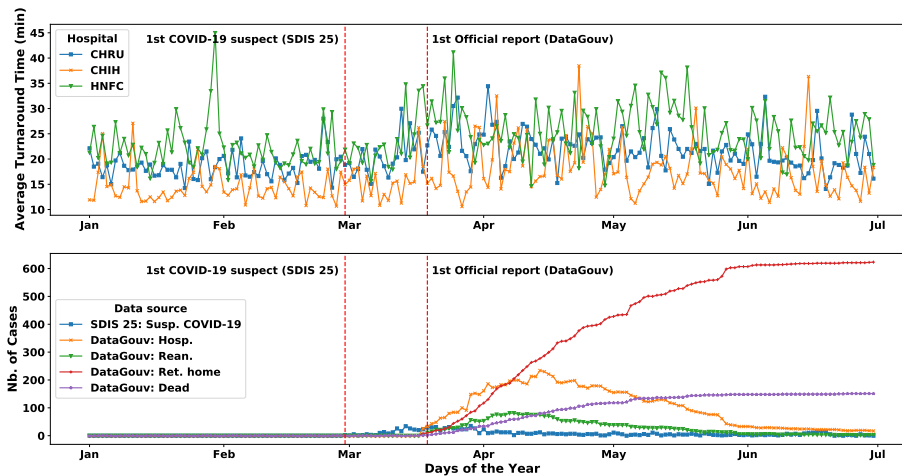


Figure 2: Statistical analysis for the AvTT per day considering the COVID-19 impact. The plot on the top illustrates statistics on the average turnaround time in minutes per day for 2020 and each hospital. The plot on the bottom illustrates the COVID-19 related statistics provided by Data Gouv and the SDIS 25.

In Figure 2, while the first day with suspicious cases of COVID-19 is 29/02/2020 reported by the SDIS 25, the date with official cases is almost three weeks later, on 18/03/2020. However, one can notice that as soon as the SDIS 25 starts to take a significant number of people with the symptoms of the disease to

hospitals in March (after 10/03/2020), the AvTT per day starts to increase for CHRU and HNFC. This increment is more remarkable, for the three hospitals, after the first official cases have been reported. Furthermore, the peak period of hospitalized cases occurred approximately from the middle to the end of April. In such conditions, there were more hospitalized patients in which some of them started to need more and even intensive care as the number of cases in reanimation indicates. When comparing the AvTT distributions of the 3 hospitals and the number of discharged cases (Ret. home), we find an inverse correlation, given that as hospital resources are freed up, AvTTs decrease. The opposite case can be found when comparing the AvTT of the 3 hospitals and the number of hospitalized cases (Hosp.), where the 4 distributions follow a similar pattern, indicating that there is a positive correlation between the variables, i.e., as the number of hospitalized patients increased, the hospitals became saturated and their response time to receive new patients was affected. This generated larger TTs for the ambulances, leaving them unavailable to attend other interventions and limiting the resources of the fire brigade. In short, a chain effect on the availability of resources.

To complement Figure 2, Table 2 exhibits the AvTT per hospital and month for 2020 highlighting in italic font the highest value per hospital. Comparing the highest AvTT and the lowest one per hospital, considerable increments were noticed. For instance, hospital CHRU presented about 24.96% more AvTT in April than in January, hospital CHIH present about 42.68% more AvTT in April than in February, and hospital HNFC presented about 29.13% more AvTT in May than in February. In general, for each hospital, the period pre-pandemic (January and February) presented low AvTT while months March, April, and May (during-peak-pandemic) had peak-values, which is in accordance with the number of hospitalized cases and reanimation ones in Figure 2. Subsequently, in June (post-peak-pandemic) and for each hospital, the AvTT had an intermediate value between the two aforementioned periods. This can be due to the increasing number of patients that returned home and did not need special care as the 'ret. home' curve in Figure 2 indicates. Proportionally, the number

of hospitalized and reanimation cases started to decrease, which alleviates the
 390 healthcare system.

Table 2: Data analysis, for each hospital, on the mean \pm std TT values in minutes per month during 2020.

	January	February	March	April	May	June
CHRU	18.34 \pm 7.75	18.93 \pm 8.39	21.88 \pm 10.62	22.90 \pm 9.12	21.55 \pm 8.89	20.38 \pm 9.45
CHIH	14.92 \pm 6.81	14.15 \pm 4.95	16.38 \pm 6.64	20.19 \pm 10.55	17.44 \pm 6.20	16.32 \pm 7.86
HNFC	23.29 \pm 10.63	21.11 \pm 7.22	25.15 \pm 11.27	24.24 \pm 9.02	27.26 \pm 9.92	24.38 \pm 9.77

With these elements in mind, it is now evident that the COVID-19 pandemic has affected the ambulances' TT in hospitals of Doubs-France. This increment might affect the SDIS 25 service to the population as their staff and resources would be in use for more time. That is, as a chain-like effect, this would incre-
 395 ment the number of breakdowns for the fire brigade when needing to prepare for future interventions, as presented in the work [11]. In the next Subsection 2.3, we present and discuss an analysis of the service breakdown during 2020 to validate the negative impact of the COVID-19 pandemic to SDIS 25.

2.3. Service Breakdown Analysis

400 Besides an in-depth analysis of the COVID-19 impact on ambulances' TT in hospitals of Doubs France, we also present an study on the chain-like effect that high TT had on the fire service. Following the work in [11], we applied the methodology for the breakage calculation for the months of March to June 2020, considering only the public service breakdowns of the type "Rescue People".
 405 Public service breakdowns occur when there are no adapted engines, firemen, or both in the centers. This means that no adapted armament is available for a certain period of time to respond to an incident.

Figure 3 shows the average breakage time in seconds per day for the months mentioned above comparing the years 2019 and 2020. One can notice small
 410 peaks during the months of March and June, which are periods before and after the peak of the pandemic, respectively. However, in the first week of April, there

is a peak break of almost 2 hours, which corresponds to the increased TT for ambulances in hospitals presented in Figure 2. What is more, April and May have more days with longer breakdown times. As previously analyzed, during these 2 months there were more cases of COVID-19 in the Doubs region, which reveals that there was certainly a chain reaction effect that left the fire service vulnerable, and as a consequence, a certain sector of the population as well.

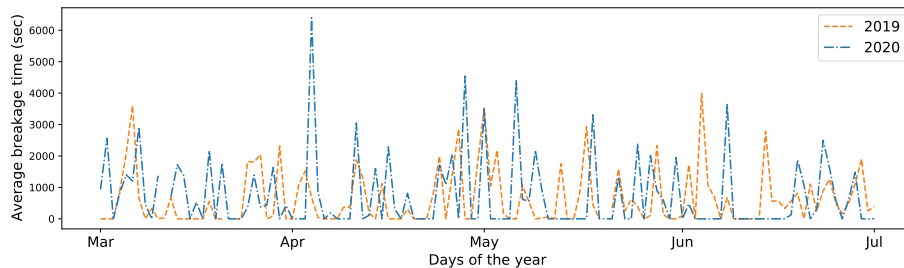


Figure 3: Average breakage time in seconds per day for 2020, considering service public breakdowns of the type “rescue people”.

In more detail, Figures 4 and 5 illustrate the causes of the breakdowns during the first semester of the years 2019 and 2020, respectively. One can notice that for January and February of both years, the results are close. However, in March 2020 there is notable increment in the number of breakdowns due to lack of engines, which is the starting month with increasing COVID-19 cases. On the other hand, April and May show a lower number of breakdowns, which is because there were fewer interventions during the lockdown period² as people spent more time in their homes. However, according to Figure 3, the breakdown time was longer, as the available adapted engine spent more time in hospitals. Also, from March to June 2020, it can be seen that there were few or no breakdowns due to the lack of firefighters. The reason behind this is that firefighters and especially volunteers were at home and showed more availability, and there were not as many simultaneous interventions as in 2019.

For this reason, this study aims at developing ML models for forecasting the TT of ambulances. Such knowledge is paramount for EMSs, as they could better prepare themselves for future interventions and avoid breakdowns on the

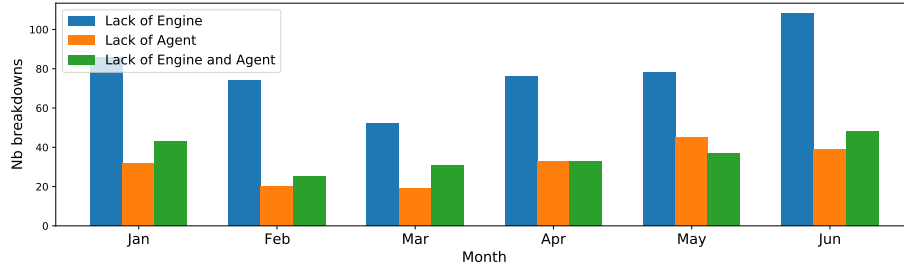


Figure 4: Causes of public service breakdowns of the type “rescue people” by month in 2019.

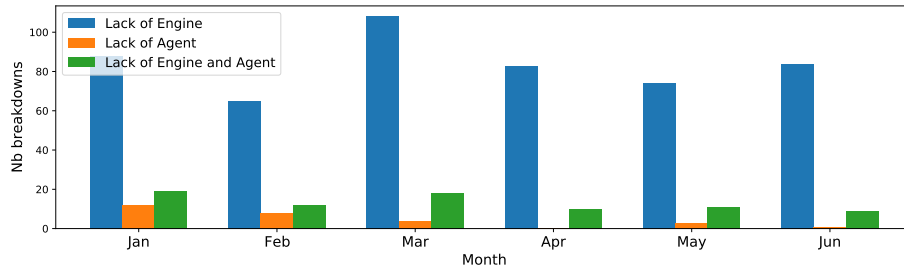


Figure 5: Causes of public service breakdowns of the type “rescue people” by month in 2020.

service as previously analyzed. Such data-driven systems should be of high
 435 confidence in order to adequately assist in decision-making solutions in real-life.
 Moreover, these systems should also be robust enough to abnormal situations
 such as natural disasters or even the current COVID-19 pandemic.

2.4. Proposed Methodology

2.4.1. Overview

440 The present research proposes a new methodology based on ML techniques
 for predicting the TT of each ambulance related to a hospital. The approach is
 composed of two time series models (i.e., regular and irregular). Figure 6 de-
 scribes the interaction of both models that are created in two stages as explained
 in the following:

445 *Stage 1: Predicting the average turnaround time per hour at each hospital.* To
 predict the TT for an ambulance, a valuable input would be the approximate
 waiting time that patients may be having in a hospital. However, we do not

have internal hospital data to make this prediction. Also, we may consider as
entry the AvTT in a given hour for ambulances from different public and private
450 organizations. However, we do not have a record of all of them or their flow.
But, what we do have is the history of ambulances arrivals over the years of
the fire department of Doubs. So, from this history we generated a regularly
spaced time series per hour, which is the AvTT-DS. In this way, it is possible
to capture trends and seasonality over time, for example: over the years the
455 AvTT per hour has been increasing, during the day the AvTT is higher than
at night, etc. Then, we developed a predictive model of AvTT for the next
hour, considering external variables from External-DS, that can influence long
waiting periods.

Stage 2: Predicting the turnaround time for an ambulance. This stage involves
460 the construction of an irregularly spaced time series model, using the Arr-DS
generated from the arrival history, it contains temporal characteristics and the
type of intervention. Likewise, we included the External-DS and the AvTT
predicted in the first stage. The objective is to forecast the TT of a specific
ambulance, that is, at the moment an ambulance warns that it will go to the
465 hospital, we make the prediction to know its approximate TT in a certain hos-
pital. In this way, firefighters will be able to establish better strategies with
their resources.

To evaluate the proposed methodology, the months April, May, and June
2020 were selected as the *testing set* for each hospital, since these are the months
470 with a higher number of COVID-19 cases (peak-pandemic period), as previously
analyzed in Subsection 2.2. For CHRU and CHIH hospitals the training set
starts from January 2015 until March 2020, and for HNFC the training set is
from February 2017 to March 2020.

During the hyperparameter search process, which is developed independently
475 for each model (AvTT and TT), each iteration tests a different configuration
throughout the training cycle. Since we use time series models, in which there
is a dependency in prior time steps that allows recognizing the increase in TT
over hours and days, the cross-validation-based training process for both models

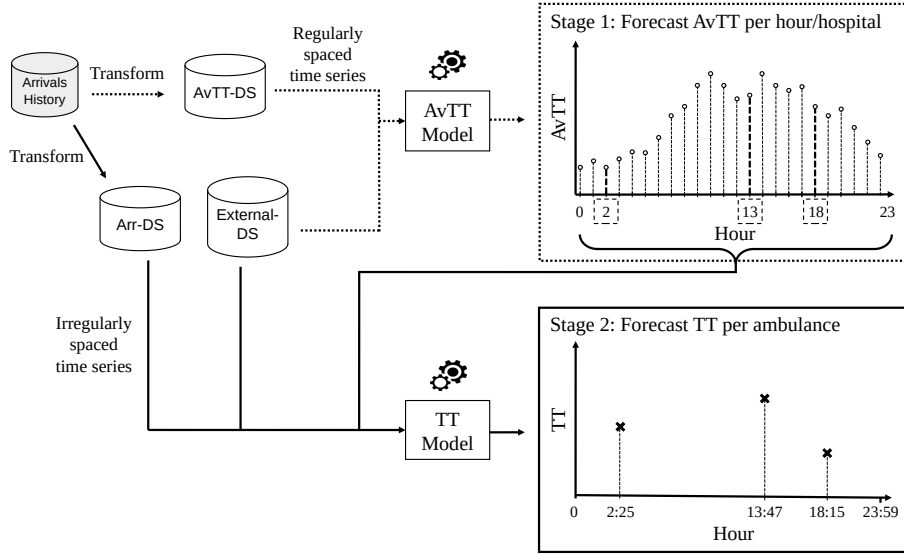


Figure 6: The proposed methodology is based on ML techniques to predict the TTT of an ambulance in a given hour and hospital. The first stage (dashed line flow) is based on a multivariate model of regularly spaced time series to predict the AvTT of the next hour in each hospital. The second stage (solid line flow) is a multivariate irregularly spaced time series model, which considers temporal variables of each ambulance departure, type of intervention, external variables and the previously predicted AvTT as inputs. For example, to predict the TT of ambulances departing to a hospital at 2:25h, 13:47h and 18:15h, the predicted AvTTs for 2h, 13h and 18h (highlighted) are used, respectively.

was a rolling-origin evaluation [32, 46] (a.k.a. forward chaining). This is a
 480 more realistic approach to re-train time series models with data that becomes
 available each time (e.g., hour or day) for further predictions. More specifically,
 our models are first fitted with the training set (with data until March 2020)
 for each hospital, making predictions *one-step-ahead* for all the hours of a day
 with the AvTT model, and per ambulance departure with the TT model for
 485 a day. This way, for all days in the testing set (April, May, and June 2020),
 the training set is expanded to include all the known values of each day and
 the process is repeated, i.e., the training set is updated day by day. Besides,
 all the predictions of each model (i.e., the three months selected as the testing
 set) in an iteration are stored, and in the end, we compute the models' metrics.

490 For the hyperparameter tuning process, we used the root mean square error as the guiding metric. Thus, a new hyperparameter configuration is generated and the search process is repeated during a given number of iterations. The pre-processing and modeling of each model, AvTT and TT, are described in the following two subsections.

495 2.4.2. AvTT model: pre-processing and modeling

The dataset for the regularly spaced time series model, that will forecast the AvTT per hour, was organized as follows:

- We took the AvTT-DS and for each sample we added 3 moving averages of AvTT as features, using a window size of 2 hours back.
- 500 • Also, the AvTT of the last 24 hours were added as features.
- Finally, we complete the dataset with the features of the External-DS, according to the previous hour, since in real life these are data that we can obtain previously.

The structure can be seen in Figure 7, where each line represents a sample by hour with its identifier (Hour ID), and the columns illustrate the predictors and the target, in yellow and orange, respectively. Next, all features were standardized using Scikit-Learn’s [47] StandardScaler function, except by the target, which is the AvTT per hour in minutes.

In the search for the best model that gives us more accurate results, we 510 compared the performance of various intelligent approaches such as traditional neural networks (MLP and LSTM), decision trees (LGBM), and a framework specifically oriented to time series (Prophet). Each technique is briefly described as follows:

- Light Gradient Boosted Machine (LGBM) [48]. It is a novel gradient 515 boosting decision tree algorithm, which uses the leaf-wise tree growth strategy to significantly reduce calculation speed and memory consumption.

Hour ID	Temporal		Google trends			Traffic		Weather			DataGouv		...	Windowing				Target			
	year	month	...	COVID-19	grippe	...	bisonDepart	bisonReturn	temperature	pressure	...	hosp.	rea.	Avg1	...	AvTT-24h	...	AvTT-1h	AvTT
...
100	2015	1	...	0	52	...	1	3	0	99390	...	0	0	15.4	...	15	...	10	25
101	2015	1	...	0	42	...	1	2	1	99323	...	0	0	12.4	...	20	...	15	18
...
47452	2020	3	...	12	93	...	4	2	12	99380	...	48	21	20.2	...	18	...	25	30
...

Figure 7: Illustrated example of the input data structure used for the AvTT model. In gray the hour identifier (Hour ID), in yellow the explanatory variables such as time, traffic, weather, statistics of COVID-19 cases from Data Gouv, trends in Google, number of assisted interventions, AvTT of the previous 24 hours, etc. And in orange, the AvTT to predict.

- 520
 • Multilayer Perceptron (MLP). It is an artificial neural network of the feed-forward type, since the information flows through several neurons organized in interconnected layers classified as: input, intermediate (which can be more than one) and output [49]. In this study, we used the implementation developed by Scikit-learn library, where the activation function considered was Rectified Linear Unit (ReLU), and the optimization algorithm for learning information was Adaptive Moment Estimation (Adam) [50].
- 525
 • Long Short-Term Memory (LSTM) [51]. It is a type of recurring neural network that overcomes the vanishing gradient problem. Inside its cell memory unit, the learning process is controlled by three gates: input, forget and output, which give it the ability to forget part of the previous memory and add new information. In this study, we used the Keras
 530
 library [52] to build our own architecture.
- Prophet [53]. It is a forecasting tool for time series data, where trends are fit with a certain seasonality, depending on the additive or multiplicative model selected. Furthermore, it allows the addition of changepoints such as holidays, and is robust enough to missing data.

535
 To navigate the hyperparameter space and pick the most optimal set, it was used the HYPEROPT library [54] with 50 iterations for each technique

described previously. This library is based on Bayesian optimization and the selected logic was the algorithm Tree Parzen Estimator suggest (tpe.suggest), which instead of modeling the probability of an observation given a new configuration, models two density functions according to a defined percentile, to finally estimate the expected improvement of a given run. The root mean square error, described in Subsection 2.6, was used as a metric to guide the search for the best configuration per technique.

2.4.3. TT model: pre-processing and modeling

The dataset for the irregularly spaced time series model, that will predict the TT for each ambulance, was constructed as follows:

- We took the Arr-DS and added the features of the External-DS, according to a previous hour, since these features come from sources that export their data online and by hour.
- Since the AvTT model runs earlier, it will give us the average prediction for the current hour, in which an ambulance is on the scene and reports that it will go to the hospital. In this way, the predicted AvTT will be used as a predictor when we forecast the real turnaround time of a specific ambulance. In addition, the AvTT per hour of the last 48 hours were added in order to improve the model's ability to recognize daily seasonality.

The structure can be seen in Figure 8, where each line represents an ambulance departure to the hospital with its identifier (Departure ID), and the columns in yellow and orange exemplify the predictors and the target, respectively. Next, all features were standardized using Scikit-Learn's StandardScaler function, except for the three variables that describe the type of intervention, which are categorical and were encoded using One-Hot-Encoding (OHE) function from Scikit-Learn library too. The target, which is the TT of an ambulance, was kept in its original format (minutes).

Departure ID	Type of intervention			AvTT			Temporal			Google trends			Traffic		Weather			DataGouv	...	Target			
	Level1	Level2	Level3	AvTT Predicted	AvTT-48h	...	AvTT-1h	year	month	...	COVID-19	grippe	...	bisonDepart	bisonReturn	temperature	pressure	...	hosp.	rea.	TT
...
200	1	2	4	25	15	...	10	2015	1	...	0	52	...	1	3	0	99370	...	0	0	22
201	1	3	4	15	20	...	15	2015	1	...	0	35	...	1	2	1	99325	...	0	0	14
...
78700	2	3	3	30	18	...	25	2020	3	...	12	100	...	4	2	9	99390	...	52	32	28
...

Figure 8: Illustrated example of the input data structure used for the TT model. In gray the depart identifier, in yellow the explanatory variables such as the type of intervention, the predicted AvTT for the present hour and the AvTT of the last 48h, time variables, Google trends, traffic, weather, etc. And in orange, the TT of the ambulance to predict.

565 To build the time series models, it was chosen the LGBM technique, because of its robustness and execution speed. To find the best set of hyperparameters, we performed 100 iterations using tpe.suggest from HYPEROPT technique. The loss function used was the root mean square error described in Subsection 2.6.

2.5. Baseline models

570 Two straightforward baseline models are used to compare the effectiveness of the proposed methodology. The first baseline is to compare with the multivariate regularly spaced time series model from Subsection 2.4.2. The second baseline is to compare with the multivariate irregularly spaced time series model from Subsection 2.4.3.

575 2.5.1. Baseline: AvTT per hour

The baseline model to predict the AvTT per hour (BSAvTT) uses the Arr-DS, where the TT of ambulances until the previous day are averaged by hour, and these values are used as a prediction for each hour during the current day. At the end of each day, if there were ambulances transporting victims to the hospital, the average for each hour is updated.

2.5.2. Baseline: TT for an ambulance

The baseline model for predicting the TT of an ambulance at a given hour and hospital (BSTT) considers adding the External-DS features, according to

a previous hour, to each ambulance in the Arr-DS. In this solution, we do not
585 consider as features any AvTT per hour of previous or predicted hours.

Similar to the standardization described in Subsection 2.4.3, predictors were transformed into features using the StandardScaler and One-Hot-Encoder from Scikit-Learn’s library. The target kept its original value in minutes. The technique used for the modeling was LGBM and for the optimization of hyperpa-
590 rameters was HYPEROPT with its tpe.suggest algorithm, 100 iterations, and the root mean square error as the loss function described in Subection 2.6

2.6. Metrics

Let y_i be the real output, \hat{y}_i be the predicted output, and n be the total number of samples, for $i \in [1, n]$. In this paper, the models were evaluated using
595 the metrics described in following.

- Mean absolute error (MAE): MAE measures the averaged absolute difference between real and predicted values and is calculated as: $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$.
- Root mean squared error (RMSE): RMSE measures the square root
600 average of the squares of the errors and is calculated as: $RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}$.
- Accuracy with margin of error ± 10 (ACC10): ACC10 is the ratio of number of correct predictions, with a maximum margin of error ± 10 , to the total number of input samples. Let $f(y_i, \hat{y}_i)$ be a function that $f(y_i, \hat{y}_i) = 1$
605 if $\hat{y}_i \in \{y_i \pm 1, y_i \pm 2, \dots, y_i \pm 10\}$ and 0 otherwise. ACC10 is calculated as: $ACC10 = \frac{1}{n} \sum_{i=1}^n f(y_i, \hat{y}_i) \cdot 100(\%)$. For example, if $y_i = 35$ and $\hat{y}_i = 28$ or $\hat{y}_i = 41$, $f(y_i, \hat{y}_i) = 1$ in both cases.

3. Results

In this section, it is first analyzed the results for the time series model to
610 predict AvTT per hour and hospital (Subsection 3.1). Further, we analyze

results for the main goal of this work, i.e., to predict the TT for each ambulance (Subsection 3.2) transporting victims to hospitals.

3.1. Forecasting the average turnaround time of ambulances per hour

To predict the AvTT of ambulances per hour and hospital during April, May, and June 2020, respectively, the baseline BSAvTT and four ML-based models were evaluated, namely LGBM, MLP, LSTM, and Prophet. Table 3 presents the metrics, discussed in Subsection 2.6, for each model and hospital, where the best results are highlighted in italic font. Both MAE and RMSE metrics express model prediction error in units of the variable of interest AvTT (minutes).

In Appendix A, Table A.5 indicates the range of each hyperparameter we considered in the Bayesian optimization, as well as the best configurations used to train and evaluate the models. The search space for each technique was established with previous empirical experiments that helped to limit the search and define the most influential hyperparameters. The latter are shown in the table for each technique while those that do not appear keep their default values. In the case of LGBM, the maximum number of boosted trees was 1000 with a maximum depth of 12, and subsets of samples and features were greater than 50% of the complete dataset. In the case of neural networks such as MLP, 2 dense layers were established, where the number of neurons varied for each layer. Unlike LSTM, where the number of layers and neurons were initially defined, and the variations occurred at the batch size and learning rate level. In the case of Prophet, the default number of n.changepoints is 25, and it was modified to vary up to 100 in order to recognize the changes in the trends of the TTs over hours, days, weeks, months and years, Similarly, varying the seasonality_prior_scale allows us to experiment with different sizes of fluctuations over time.

Moreover, to highlight the effectiveness of the techniques in some peak values, Figure 9 illustrates for each hospital and two days (48h), the comparison of each ML model to forecasting the AvTT of ambulances per hour in different days. Lastly, Figure 10 illustrates for each hospital, the time series results for the period with more breakdowns in service (last weeks of April and the first

Table 3: Prediction results for the AvTT of ambulances per hour and hospital by ML technique.

Hospital	Metric	BSAvTT	LGBM	MLP	LSTM	Prophet
CHRU	RMSE	9.70	7.36	7.78	8.77	7.70
	MAE	6.82	4.94	5.49	6.05	5.34
	ACC10	83.29%	90.16%	88.05%	84.02%	88.46%
CHIH	RMSE	9.08	3.79	4.27	4.85	4.00
	MAE	5.60	1.76	2.60	2.75	2.15
	ACC10	86.81%	97.02%	96.79%	95.28%	96.98%
HNFC	RMSE	10.08	5.97	6.77	7.88	6.66
	MAE	7.58	3.52	4.45	5.53	4.55
	ACC10	77.24%	93.09%	91.30%	86.22%	91.48%

weeks of May 2020), according to Figure 3, during the COVID-19 peak-period. This figure considers only the LGBM model, according to the best results of Table 3, and the baseline BSAvTT.

One can notice in Table 3 that all ML-based methods consistently and considerably outperform the baseline BSAvTT model. Among the four ML-based models, LGBM presented the best performance for all metrics and hospitals. Similar metric results were achieved by MLP and Prophet. Additionally, LGBM provided faster execution time than neural network-based methods (MLP and LSTM) and Prophet. For hospital CHIH, LGBM provided lower estimation error with $RMSE = 3.79$ and $MAE = 1.76$ minutes, which naturally led to more correct predictions, considering a margin of error of ± 10 minutes, and $ACC10 = 97.02\%$. These results demonstrate that it is possible to forecast the AvTT of ambulances in hospitals per hour with good accuracy for practical purposes, i.e., 97% of the time, the error is less than or equal to 10 minutes. Similarly, although not as good as for hospital CHIH, hospitals CHRU and HNFC present ACC10 metric higher than 90% with RMSE about 6 to 7 minutes.

One can notice in Figure 9 that ML-based models follow the AvTT hourly trend, as well as some peak values. In Figure 10, while LGBM presents accurate prediction results for the AvTT of ambulances for all three hospitals, BSAvTT presents poor predictions following a similar pattern for them through the days.

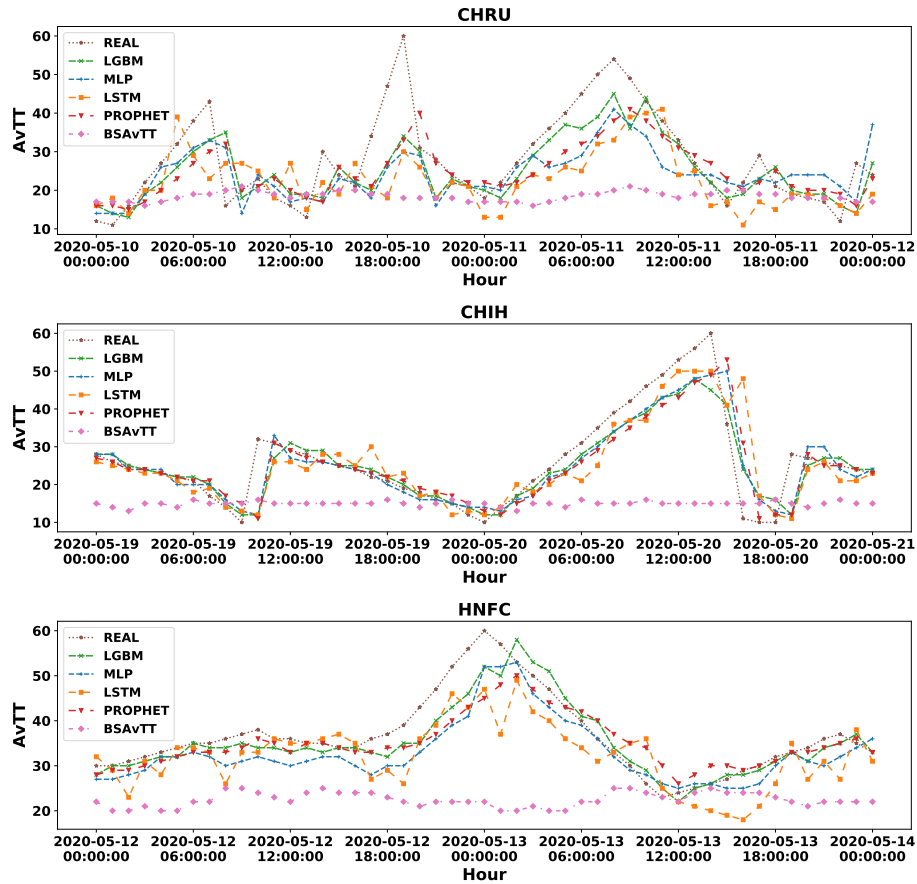


Figure 9: Comparison between the real and the predicted AvTT for ambulances per hour in each hospital, namely CHRU, CHIH, and HNFC, respectively, by each ML technique and the baseline BSAvTT.

Due to COVID-19, the TTs of ambulances increased during some periods (cf. Figure 2), and therefore, using only the average per hour resulted in poor performance.

3.2. Forecasting the turnaround time for each ambulance

665 To predict the TT of each ambulance in a given time and hospital during April, May, and June 2020, respectively, the baseline BSTT and our proposed methodology were evaluated. In both cases, LGBM is the modeling technique.

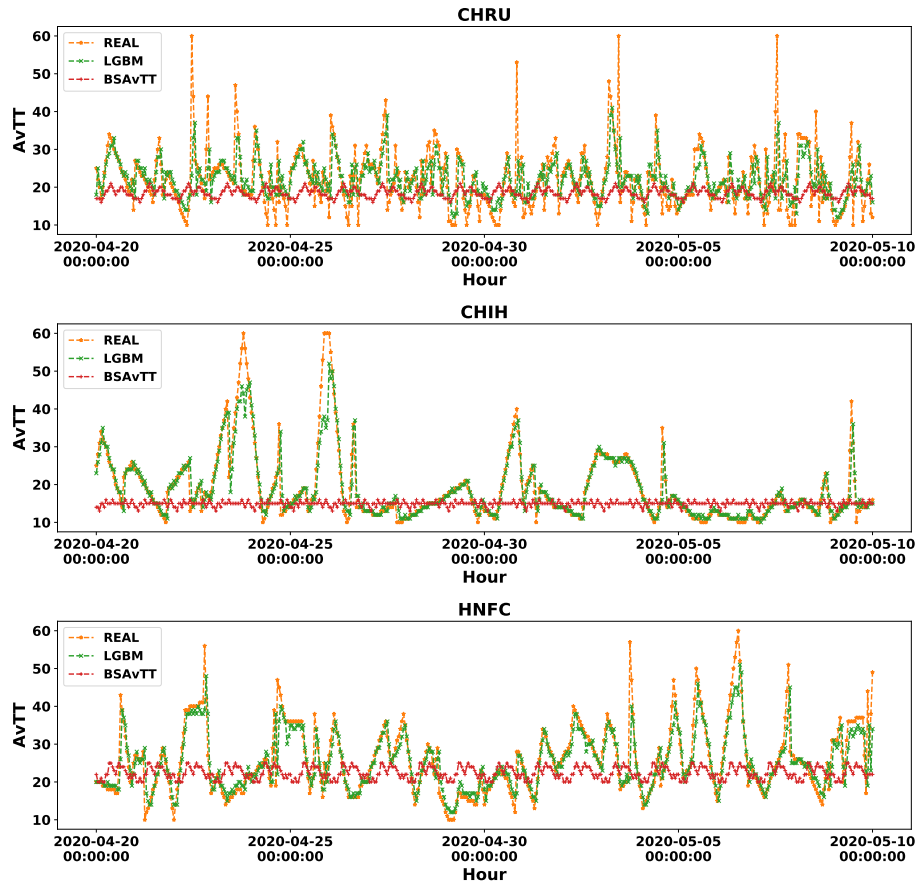


Figure 10: Comparison between the real hourly AvTT, the predictions of the best model obtained with the LGBM technique and the predictions of the baseline model BSAvTT, for each hospital (CHRU, CHIH and HNFC). It is observed that the AvTT predicted by LGBM better recognizes the patterns of the mean times and their variation through time, while the BSAvTT maintains a more constant pattern.

In our proposal, there are additional predictors such as the AvTT predicted by the first (and best) regularly spaced multivariate time series model, i.e., also using LGBM, as well as past AvTTs for recognizing average daily trends. Table 3 presents the metrics, discussed in Subsection 2.6, for each model and hospital, where the best results are highlighted in italic font. Both MAE and RMSE metrics express model prediction error in units of the variable of interest TT

(minutes). In Appendix A, Table A.6 indicates the range of each hyperparameter we considered in the Bayesian optimization, as well as the best configurations used to train and evaluate the models. The description of the LGBM search space follows the one made in section 3.1

Table 4: Prediction results for the TT using the proposed methodology and the BSTT model.

Hospital	Metric	BSTT	Proposed methodology
CHRU	RMSE	13.42	<i>12.92</i>
	MAE	8.74	<i>8.31</i>
	ACC10	73.96	<i>74.42</i>
CHIH	RMSE	13.99	<i>10.53</i>
	MAE	7.70	<i>5.37</i>
	ACC10	82.97	<i>86.63</i>
HNFC	RMSE	13.31	<i>11.70</i>
	MAE	9.58	<i>7.74</i>
	ACC10	68.02	<i>76.67</i>

Moreover, Figure 11 illustrates the comparison of each model to forecasting the ambulance’s TT in a given time and hospital, in different periods during the COVID-19 outbreak. Notice that there might be several or no ambulance in a given hour in the hospital. Also, this figure plots the ambulances’ TT according to the moment in which they reported that they were going to the hospital.

As one can notice in Table 4, the proposed methodology consistently outperforms the straightforward BSTT model. While for hospital CHRU the difference is small and favoring our methodology, for hospital CHIH and HNFC, our proposal considerably outperforms the baseline model. For instance, our proposed methodology provided a lower estimation error with $RMSE = 10.53$ and $MAE = 5.37$ minutes for hospital CHIH, which naturally led to more correct predictions by achieving $ACC10 = 86.63\%$. Similarly, hospitals CHRU and HNFC present ACC10 metric around 75% with RMSE about 12 to 13 minutes. These results demonstrate that it is possible to forecast the TT of ambulances in a given time and hospital with good accuracy for practical purposes, i.e., 86% of the time, the error is less than or equal to 10 minutes for CHIH, for example. These results are also reflected in Figure 11, where our proposed methodology

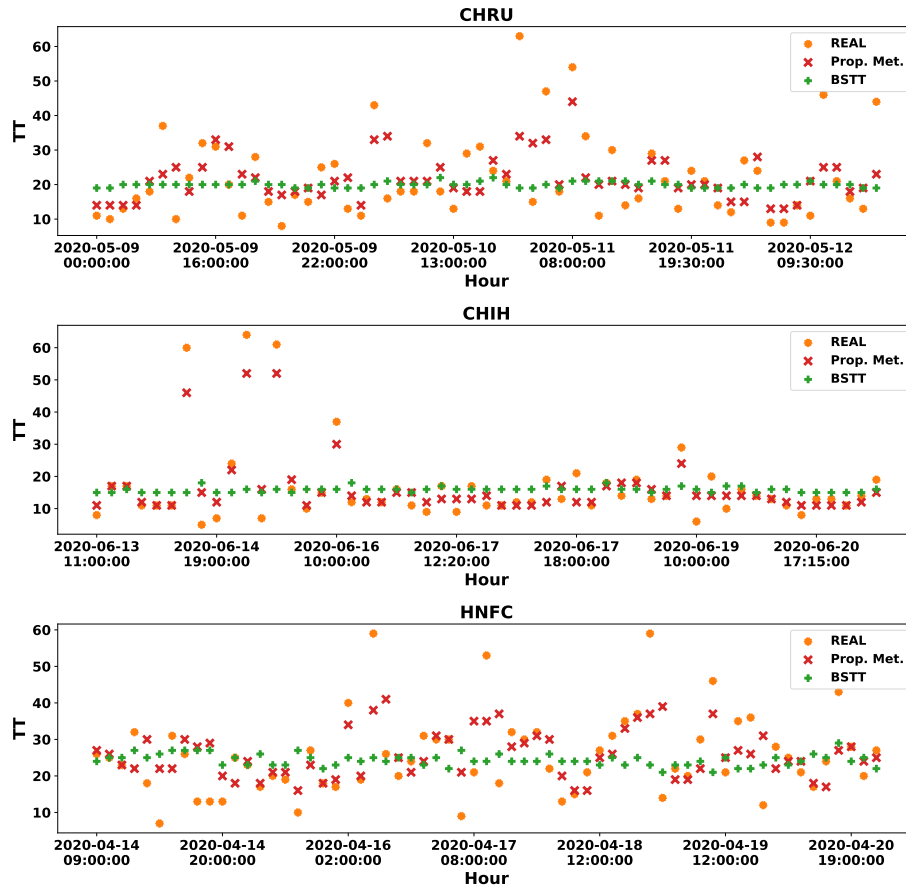


Figure 11: Comparison between the real TT of each ambulance for each hospital (CHRU, CHIH and HNFC), the predictions of the proposed methodology and the predictions of the baseline model BSTT.

695 recognizes high- and low-peak TTs of ambulances. On the other hand, BSTT tends to underestimate the TTs for many ambulances by forecasting values around a mean value.

4. Discussion

700 Extended waiting times to transfer patient(s) from ambulances to EDs (i.e., AOD) and eventually, high TTs for ambulances, may lead to numerous consequences for patients' care, financial losses, and unavailability of providing

adequate emergency medical services [7, 3, 9, 8, 4], for example. In this paper, we analyzed ambulances' TT from January 2015 to June 2020 on two hospitals (CHRU and CHIH), and from February 2017 to June 2020 on one hospital
705 (HNFC), for the SDIS 25 fire department in Doubs, France. Although one can find works in the literature investigating the TTs of ambulances [6, 34, 2, 33], we included in our analysis (Subsection 2.2) the impact of COVID-19 in the TTs of ambulances in a larger longitudinal study. Further, we analyzed the negative impact due to COVID-19, which increased the TTs of firefighters' ambulances
710 and, consequently, generated more breakdowns in services (Subsection 2.3).

The study in [34] identified a high correlation between patient acuity and TT. In our case, there are three variables related to the type of intervention, in which the latter is the severity level of the victim. Taking this into consideration and for the first time, this research proposed a ML-based methodology (cf. Figure 6)
715 to forecast the TT for each ambulance in a given time and hospital, using as a reference the forecasting of AvTT per hour for ambulances in hospitals, its type of intervention, and external variables. Indeed, forecasting the TT of an ambulance in a given time and hospital could provide valuable information for SDIS 25, and in general, for EMSs. For instance, EMSs could activate proactive
720 decision-making with the available resources and personnel in order to be able to save more lives. Further, if there are policies on ambulance diversion, EMSs may consider diverging their ambulances to other EDs whose TT is smaller to provide adequate care to the ambulances' patient(s), as well as, returning on service in less time. Besides, the predicted AvTT for ambulances in the
725 next hour(s) could provide approximate information for ambulances' TT, which hospitals and EMSs services may consider as a reference value, and use as a predictor in a second model as we propose.

These data-driven systems should be of high confidence in order to assist adequately as a decision-support tool in real-life. Moreover, these systems should
730 also be robust enough to abnormal situations such as natural disasters and pandemics. For instance, in adverse cases such as the current COVID-19 pandemic, the need for such a kind of information (ambulances' TT and AvTT) becomes

vital since healthcare systems may saturate. While our work was motivated to include the impact of the novel COVID-19 pandemic, our solutions are not
735 limited to it. Indeed, these forecasting models could help other private or public EMSs to forecast the TT of their ambulances and AvTT of hospitals, as well as to future abnormal situations.

To evaluate our proposed methodology, experiments were performed for the period of April, May, and June of 2020, which considers the peak-period of the
740 COVID-19 first wave in the Doubs Region (cf. Figure 2) and a high number of breakdowns in the SDIS 25 service (cf. Figures 3 and 5). As shown in the results, it is possible to accurately forecast the TT of each ambulance in a given time and hospital, as well as, to forecast the AvTT of ambulances per hour and hospital. On the one hand, our proposed methodology achieves ACC10
745 metrics ranging from $\sim 75\%$ (CHRU) to $\sim 87\%$ (CHIH), considering a margin of error of ± 10 minutes, which are promising accuracies for practical purposes to forecasting the TT of an ambulance. In addition, the multivariate regularly spaced time series model trained with LGBM achieves ACC10 higher than 90% for all three hospitals for predicting AvTT.

In our experiments, it is remarkable the improvement of the results with
750 the proposed methodology, for such complex and important tasks, comparing to straightforward prediction models such as the baselines BSAvTT and BSTT. On the one hand, BSAvTT is a straightforward solution that considers only the historical data of ambulances' TT and averages the TT per hour for prediction. Even though this is an intuitive solution that hospital staff and/or EMS could
755 think of applying, we demonstrate that training multivariate time series models based on ML result in much higher performance. For instance, external variables such as meteorological (bad weather may results in more traffic accidents, floods, etc), temporal (day of the week/year, holiday or not), trends (google search for
760 disease-related keywords), are variables that may lead to reasons of overcrowding in hospitals' EDs. On the other hand, BSTT is a straightforward model that one might consider applying by using only the historical data of ambulances' TT, external variables (traffic, weather, etc.), and variables related to the incident

(type, hour, etc.). However, we demonstrate that adding the predicted AvTT of
765 ambulances for the hour an ambulance is going to the hospital and past AvTTs,
leads to much higher performance for this complex task.

Finally, the present work has some key limitations that are described in
the following. We analyzed and build our models using the data of SDIS 25.
Although it may represent a sufficient amount of samples, other emergency
770 medical services in France also transport victims to hospitals' EDs. In addition,
there was no data from hospitals such as a history of all ambulances' TT, patient
flow and the number of doctors on duty, for example. These variables could be of
high importance as they are internal to hospitals, which would help ML models
to better forecast the TT of each ambulance and AvTT for ambulances.

775 **5. Conclusion and Future Work**

This paper presents a first study on the impact of the current COVID-
19 pandemic on the firefighter ambulances' TT on hospitals in the region of
Doubs-France. A significant increase in the AvTT per day has been identified
in 2020 as soon as the number of official and suspected cases of the disease
780 began to rise. Additionally, in comparison with previous years, this increment
is not normal, which strengthens the claim that the AvTT per day increased
due to the COVID-19 pandemic. A direct and negative chain-like effect is the
increment of breakdowns on the firefighters' service in 2020 due to the lack of
ambulances and personnel in the centers of the SDIS 25. Therefore, it is vital to
785 have some data-driven system to forecast the TT that will have an ambulance in
a certain hospital, as we propose in this paper. This prediction could be made
at the moment when the personnel report that they will go to the hospital. This
would help fire brigades, and in a global context, EMSs to activate proactive
decision-making with the available resources in order to allow saving more lives.

790 For future work, we will extend our analyses and forecasting to the sec-
ond semester of 2020, which includes the current second wave of COVID-19 in
France. Also, we aim to improve the proposed methodology by adding new

features and analyzing their influence (i.e., feature selection), and to test with other ML techniques. Finally, there is also another direction for further research, and that is about finding the nearest hospital with a shorter TT, considering
795 internal hospital data.

Acknowledgment

This work was funded by SDIS du Doubs, with the support of the French Ministry of Higher Education and Research (managed by the National Association
800 of Research and Technology (ANRT) for the CIFRE thesis (N 2019/0372), by the EIPHI-BFC Graduate School (contract “ANR-17-EURE-0002”), and by the Region of Bourgogne Franche-Comté CADRAN Project. Computations have been performed on the supercomputer facilities of “Mésocentre de Calcul de Franche-Comté”.

805 References

- [1] D. C. Cone, P. M. Middleton, S. M. Pour, Analysis and impact of delays in ambulance to emergency department handovers, *Emergency Medicine Australasia* 24 (5) (2012) 525–533. <https://doi.org/10.1111/j.1742-6723.2012.01589.x>.
- 810 [2] Y. J. Lee, S. D. Shin, E. J. Lee, J. S. Cho, W. C. Cha, Emergency department overcrowding and ambulance turnaround time, *PLOS ONE* 10 (6) (2015) e0130758. <https://doi.org/10.1371/journal.pone.0130758>.
- [3] E. Segal, V. Verter, A. Colacone, M. Afilalo, The in-hospital interval: A description of EMT time spent in the emergency department, *Prehospital
815 Emergency Care* 10 (3) (2006) 378–382. <https://doi.org/10.1080/10903120600725884>.
- [4] M. Eckstein, S. M. Isaacs, C. M. Slovis, B. J. Kaufman, J. R. Loflin, R. E. O’Connor, P. E. Pepe, Facilitating EMS turnaround intervals at hospitals

- in the face of receiving facility overcrowding, *Prehospital Emergency Care* 9 (3) (2005) 267–275. <https://doi.org/10.1080/10903120590962102>.
820
- [5] A. J. Carter, J. Overton, M. Terashima, D. C. Cone, Can emergency medical services use turnaround time as a proxy for measuring ambulance offload time?, *The Journal of Emergency Medicine* 47 (1) (2014) 30–35. <https://doi.org/10.1016/j.jemermed.2013.08.109>.
- [6] D. C. Cone, S. J. Davidson, Q. Nguyen, A time-motion study of the emergency medical services turnaround interval, *Annals of Emergency Medicine* 31 (2) (1998) 241–246. [https://doi.org/10.1016/s0196-0644\(98\)70314-2](https://doi.org/10.1016/s0196-0644(98)70314-2).
825
- [7] M. Li, P. Vanberkel, A. J. E. Carter, A review on ambulance offload delay literature, *Health Care Management Science* 22 (4) (2018) 658–675. <https://doi.org/10.1007/s10729-018-9450-x>.
830
- [8] M. Eckstein, L. S. Chan, The effect of emergency department crowding on paramedic ambulance availability, *Annals of Emergency Medicine* 43 (1) (2004) 100–105. [https://doi.org/10.1016/s0196-0644\(03\)00747-9](https://doi.org/10.1016/s0196-0644(03)00747-9).
- [9] M. Perry, D. Carter, The ethics of ambulance ramping, *Emergency Medicine Australasia* 29 (1) (2016) 116–118. <https://doi.org/10.1111/1742-6723.12625>.
835
- [10] H. H. Arcolezi, J.-F. Couchot, S. Cerna, C. Guyeux, G. Royer, B. A. Bouna, X. Xiao, Forecasting the number of firefighter interventions per region with local-differential-privacy-based data, *Computers & Security* 96 (2020) 101888. <https://doi.org/10.1016/j.cose.2020.101888>.
840
- [11] S. Cerna, C. Guyeux, G. Royer, C. Chevallier, G. Plumerel, Predicting fire brigades operational breakdowns: A real case study, *Mathematics* 8 (8) (2020) 1383. <https://doi.org/10.3390/math8081383>.

- 845 [12] World-Health-Organization, Novel coronavirus – china, available online:
<https://www.who.int/csr/don/12-january-2020-novel-coronaviruses-china/en/> (accessed on 07 September 2020).
- [13] C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, A novel coronavirus outbreak of global health concern, *The Lancet* 395 (10223) (2020) 470–473.
850 [https://doi.org/10.1016/s0140-6736\(20\)30185-9](https://doi.org/10.1016/s0140-6736(20)30185-9).
- [14] C. Sohrabi, Z. Alsafi, N. O’Neill, M. Khan, A. Kerwan, A. Al-Jabir, C. Iosifidis, R. Agha, World health organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19), *International Journal of Surgery* 76 (2020) 71–76. <https://doi.org/10.1016/j.ijssu.2020.02.034>.
855
- [15] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia, *New England Journal of Medicine* 382 (13) (2020) 1199–1207. [10.1056/nejmoa2001316](https://doi.org/10.1056/nejmoa2001316).
860
- [16] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected with 2019 novel coronavirus in wuhan, china, *The Lancet* 395 (10223) (2020) 497–506. [10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5).
- 870 [17] A. Dairi, F. Harrou, A. Zeroual, M. M. Hittawe, Y. Sun, Comparative study of machine learning methods for COVID-19 transmission forecasting, *Journal of Biomedical Informatics* 118 (2021) 103791. [10.1016/j.jbi.2021.103791](https://doi.org/10.1016/j.jbi.2021.103791).

- [18] I. Rahimi, F. Chen, A. H. Gandomi, A review on COVID-19 forecasting models, *Neural Computing and Applications* 10.1007/s00521-020-05626-8. 875
- [19] World-Health-Organization, Naming the coronavirus disease (COVID-19) and the virus that causes it, available online: [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it#:~:text=The%20International%20Committee%20on%20Taxonomy,two%20viruses%20are%20different](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it#:~:text=The%20International%20Committee%20on%20Taxonomy,two%20viruses%20are%20different) (accessed on 07 September 2020). 880
- [20] World-Health-Organization, WHO announces covid-19 outbreak a pandemic, available online: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/news/news/2020/3/who-announces-covid-19-outbreak-a-pandemic> (accessed on 07 September 2020). 885
- [21] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. P. y Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini, A. Vespignani, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak, *Science* 368 (6489) (2020) 395–400. 10.1126/science.aba9757. 890
- [22] J. T. Wu, K. Leung, G. M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in wuhan, china: a modelling study, *The Lancet* 395 (10225) (2020) 689–697. 10.1016/s0140-6736(20)30260-9. 895
- [23] C. Anastassopoulou, L. Russo, A. Tsakris, C. Siettos, Data-based analysis, modelling and forecasting of the COVID-19 outbreak, *PLOS ONE* 15 (3) (2020) e0230405. 10.1371/journal.pone.0230405. 900
- [24] G. Grasselli, A. Pesenti, M. Cecconi, Critical care utilization for the

COVID-19 outbreak in lombardy, italy, JAMA 323 (16) (2020) 1545.
10.1001/jama.2020.4031.

- [25] S. B. Stoecklin, P. Rolland, Y. Silue, A. Mailles, C. Campese, A. Simondon,
905 M. Mechain, L. Meurice, M. Nguyen, C. Bassi, E. Yamani, S. Behillil,
S. Ismael, D. Nguyen, D. Malvy, F. X. Lescure, S. Georges, C. Lazarus,
A. Tabai, M. Stempfelet, V. Enouf, B. Coignard, D. L.-B. and, First cases of
coronavirus disease 2019 (COVID-19) in france: surveillance, investigations
and control measures, january 2020, Eurosurveillance 25 (6). <https://doi.org/10.2807/1560-7917.es.2020.25.6.2000094>.
910
- [26] D. G. FR, Données hospitalières relatives à l'épidémie de covid-19, available
online: <https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/> (accessed on 07 September
2020).
- 915 [27] W. Carroll, V. Strenger, E. Eber, F. Porcaro, R. Cutrera, D. Fitzgerald,
I. Balfour-Lynn, European and united kingdom COVID-19 pandemic experience:
The same but different, Paediatric Respiratory Reviews 35 (2020)
50–56. <https://doi.org/10.1016/j.prrv.2020.06.012>.
- [28] K. Goniewicz, A. Khorram-Manesh, A. J. Hertelendy, M. Goniewicz,
920 K. Naylor, F. M. Burkle, Current response and management decisions of
the european union to the COVID-19 outbreak: A review, Sustainability
12 (9) (2020) 3838. <https://doi.org/10.3390/su12093838>.
- [29] D. Hendricks, With hospitals full of coronavirus patients, ambulances wait
hours for beds to become available, <https://www.krgv.com/news/with-hospitals-full-of-coronavirus-patients-ambulances-wait-hours-for-beds-to-become-available/>, (Online; accessed 01 November 2020).
925
- [30] T. Homery, Quimper. l'hôpital, le calvaire des ambulanciers, <https://www.ouest-france.fr/bretagne/quimper-29000/quimper-1-hopital-le-calvaire-des-ambulanciers-6857042>, (Online; accessed 01 November
930 2020).

- [31] A. Dow, M. Cunningham, Patient deaths after wait in ambulances lead to hospital probe, <https://www.theage.com.au/national/victoria/patient-deaths-after-wait-in-ambulances-lead-to-hospital-probe-20200701-p557xg.html>, (Online; accessed 01 November 2020).
- 935 [32] R. J. Hyndman, G. Athanasopoulos, *Forecasting: principles and practice*, OTexts, 2018.
- [33] A. Clarey, M. Allen, S. Brace-McDonnell, M. W. Cooke, Ambulance handovers: can a dedicated ED nurse solve the delay in ambulance turnaround times?, *Emergency Medicine Journal* 31 (5) (2013) 419–420. <https://doi.org/10.1136/emermed-2012-202258>.
- 940 [34] S. Vandeventer, J. R. Studnek, J. S. Garrett, S. R. Ward, K. Staley, T. Blackwell, The association between ambulance hospital turnaround times and patient acuity, destination hospital, and time of day, *Prehospital Emergency Care* 15 (3) (2011) 366–370. <https://doi.org/10.3109/10903127.2011.561412>.
- 945 [35] R. Ding, M. L. McCarthy, J. S. Desmond, J. S. Lee, D. Aronsky, S. L. Zeger, Characterizing waiting room time, treatment time, and boarding time in the emergency department using quantile regression, *Academic Emergency Medicine* 17 (8) (2010) 813–823. <https://doi.org/10.1111/j.1553-2712.2010.00812.x>.
- 950 [36] Y. Sun, K. L. Teow, B. H. Heng, C. K. Ooi, S. Y. Tay, Real-time prediction of waiting time in the emergency department, using quantile regression, *Annals of Emergency Medicine* 60 (3) (2012) 299–308. <https://doi.org/10.1016/j.annemergmed.2012.03.011>.
- 955 [37] Y.-H. Kuo, N. B. Chan, J. M. Leung, H. Meng, A. M.-C. So, K. K. Tsoi, C. A. Graham, An integrated approach of machine learning and systems thinking for waiting time prediction in an emergency department, *International Journal of Medical Informatics* 139 (2020) 104143. <https://doi.org/10.1016/j.ijmedinf.2020.104143>.

- 960 [38] H.-Y. Lo, C.-H. Chaou, Y.-C. Chang, C.-J. Ng, S.-Y. Chen, Prediction of emergency department volume and severity during a novel virus pandemic: Experience from the COVID-19 pandemic, *The American Journal of Emergency Medicine* <https://doi.org/10.1016/j.ajem.2020.07.084>.
- [39] K. ArunKumar, D. V. Kalaga, C. M. S. Kumar, G. Chilkoor, M. Kawaji, 965 T. M. Brenza, Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-regressive integrated moving average (ARIMA) and seasonal auto-regressive integrated moving average (SARIMA), *Applied Soft Computing* 103 (2021) 107161. [10.1016/j.asoc.2021.107161](https://doi.org/10.1016/j.asoc.2021.107161).
- 970 [40] M. H. D. M. Ribeiro, R. G. da Silva, V. C. Mariani, L. dos Santos Coelho, Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for brazil, *Chaos, Solitons & Fractals* 135 (2020) 109853. [10.1016/j.chaos.2020.109853](https://doi.org/10.1016/j.chaos.2020.109853).
- 975 [41] J. Devaraj, R. M. Elavarasan, R. Pugazhendhi, G. Shafiullah, S. Ganesan, A. K. Jeysree, I. A. Khan, E. Hossain, Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant?, *Results in Physics* 21 (2021) 103817. [10.1016/j.rinp.2021.103817](https://doi.org/10.1016/j.rinp.2021.103817).
- [42] R. G. da Silva, M. H. D. M. Ribeiro, V. C. Mariani, L. dos Santos Coelho, 980 Forecasting brazilian and american COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables, *Chaos, Solitons & Fractals* 139 (2020) 110027. [10.1016/j.chaos.2020.110027](https://doi.org/10.1016/j.chaos.2020.110027).
- [43] K. Pirklbauer, R. D. Findling, Predicting the category of fire department operations, in: *Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services*, ACM, 2019. 985 <https://doi.org/10.1145/3366030.3366113>.
- [44] Bison-Futé, Les prévisions de trafic, available online: <https://www.bison-fute.gouv.fr> (accessed on 02 July 2020).

- [45] Météo-France, Données publiques, available online: https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=90&id_rubrique=32 (accessed on 02 July 2020).
- [46] L. J. Tashman, Out-of-sample tests of forecasting accuracy: an analysis and review, *International Journal of Forecasting* 16 (4) (2000) 437–450. [10.1016/s0169-2070\(00\)00065-0](https://doi.org/10.1016/s0169-2070(00)00065-0).
- [47] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.
- [48] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, Lightgbm: A highly efficient gradient boosting decision tree, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), *Advances in Neural Information Processing Systems* 30, Curran Associates, Inc., 2017, pp. 3146–3154.
- [49] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016, <http://www.deeplearningbook.org>.
- [50] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization (2014). [arXiv:1412.6980](https://arxiv.org/abs/1412.6980).
- [51] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Computation* 9 (8) (1997) 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [52] F. Chollet, et al., Keras, <https://keras.io> (2015).
- [53] S. J. Taylor, B. Letham, Forecasting at scale, *PeerJ Preprints* 5:e3190v2. <https://doi.org/10.7287/peerj.preprints.3190v2>.

- 1015 [54] J. Bergstra, D. Yamins, D. D. Cox, Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures, in: Proceedings of the 30th International Conference on International Conference on Machine Learning, ICML'13, JMLR, 2013, p. I-115-I-123.

Appendix A. Settings used during modeling

Table A.5: Search space for hyperparameters by technique and the best configuration obtained for predicting AvTT per hour for each hospital.

Technique	Search space	Best configuration		
		CHRU	CHIH	HNFC
LGBM	max_depth: [1-12]	6	12	5
	n_estimators: [50-1000]	143	808	768
	num_leaves: [31-100]	40	61	45
	learning_rate: [0.001-1]	0.8498	0.0103	0.0094
	subsample: [0.5-1]	0.99	0.92	0.5
	colsample_bytree: [0.5-1]	0.94	0.6	0.63
MLP	Dense layers: 2	2	2	2
	nb_neurons: [100-500]	(100, 302)	(125, 325)	(362, 562)
	alpha: [0.00001-0.01]	0.000029	0.003674	0.000034
	learning_rate_init: [0.0001-0.1]	0.0355	0.003611	0.000543
	max_iter: [50-200]	154	86	177
	tol: [0.00001-0.01]	0.003731	0.002888	0.000661
	momentum: [0.00001-0.01]	0.0073	0.006057	0.00479
	Early stopping: 20	20	20	20
LSTM	LSTM layers and neurons: 1, (110)	1, (110)	1, (110)	1, (110)
	Dense layers and neurons: 2, (128, 1)	2, (128, 1)	2, (128, 1)	2, (128, 1)
	Activation function: ReLU	ReLU	ReLU	ReLU
	Dropout: 0.5	0.5	0.5	0.5
	Loss function: 'mse'	'mse'	'mse'	'mse'
	Optimizer: Adam	Adam	Adam	Adam
	Early stopping: 15	15	15	15
	Max. epochs: 100	100	100	100
	Batch size: [40-250]	93	142	112
	Learning rate: [0.005-0.01]	0.00841	0.00595	0.00894
Prophet	n_changepoints: [20-100]	45	75	35
	seasonality_prior_scale: [0-50]	36.82	18.17	12.13
	holidays_prior_scale: [0-50]	23.47	22.42	34.77

Table A.6: Search space for hyperparameters in LGBM and the best configuration obtained according to the method applied for predicting the TTs of each ambulance.

Method	LGBM search space	Best configuration		
		CHRU	CHIH	HNFC
BSTT	max_depth: [1-12]	2	1	1
	n_estimators: [50-1000]	978	820	678
	num_leaves: [31-100]	98	59	99
	learning_rate: [0.001-1]	0.0034	0.0099	0.0157
	subsample: [0.5-1]	0.78	0.68	0.55
	colsample_bytree: [0.5-1]	0.86	0.89	0.92
Proposed methodology	max_depth: [1-12]	6	2	9
	n_estimators: [50-1000]	150	129	401
	num_leaves: [31-100]	94	46	75
	learning_rate: [0.001-1]	0.1888	0.0956	0.0299
	subsample: [0.5-1]	0.72	0.65	0.79
	colsample_bytree: [0.5-1]	0.73	0.76	0.5