

Article

Vital Signs Prediction for COVID-19 Patients in ICU

Ahmed Youssef Ali Amer ^{1,2}, Femke Wouters ^{3,4,5,6}, Julie Vranken ^{3,4,5,6}, Pauline Dreesen ^{3,4,5,6},
Dianne de Korte-de Boer ⁷, Frank van Rosmalen ⁸, Bas C. T. van Bussel ⁸, Valérie Smit-Fun ⁷,
Patrick Duflot ⁹, Julien Guiot ¹⁰, Iwan C. C. van der Horst ⁸, Dieter Mesotten ^{3,4,5,6}, Pieter Vandervoort ^{3,4,5,6},
Jean-Marie Aerts ² and Bart Vanrumste ^{1,*}

- ¹ E-MEDIA, STADIUS, Department of Electrical Engineering (ESAT), Campus Group T, KU Leuven, 3000 Leuven, Belgium; ahmed.youssefaliamer@kuleuven.be
- ² Measure, Model & Manage Bioresponses (M3-BIORES), Department of Biosystems, KU Leuven, 3000 Leuven, Belgium; jean-marie.aerts@kuleuven.be
- ³ Limburg Clinical Research Center/Mobile Health Unit, Faculty of Medicine and Life Sciences, Hasselt University, 3500 Hasselt, Belgium; femke.wouters@uhasselt.be (F.W.); julie.vranken@uhasselt.be (J.V.); pauline.dreesen@uhasselt.be (P.D.); dieter.mesotten@zol.be (D.M.); pieter@vandervoort.mobi (P.V.)
- ⁴ Limburg Clinical Research Center/Mobile Health Unit, Ziekenhuis Oost-Limburg, 3600 Genk, Belgium
- ⁵ Department of Anesthesiology, Ziekenhuis Oost-Limburg, 3600 Genk, Belgium
- ⁶ Department of Cardiology and Future Health, Ziekenhuis Oost-Limburg, 3600 Genk, Belgium
- ⁷ Department of Anesthesiology and Pain Management, Maastricht University Medical Centre+, 6229 HX Maastricht, The Netherlands; dianne.de.korte@mumc.nl (D.d.K.-d.B.); v.smit.fun@mumc.nl (V.S.-F.)
- ⁸ Department of Intensive Care, Maastricht University Medical Centre+, 6229 HX Maastricht, The Netherlands; frank.van.rosmalen@mumc.nl (F.v.R.); bas.van.bussel@mumc.nl (B.C.T.v.B.); iwan.vander.horst@mumc.nl (I.C.C.v.d.H.)
- ⁹ Service des Applications Informatiques, Centre Hospitalier Universitaire de Liège—CHU, 4000 Liège, Belgium; pduflot@chuliege.be
- ¹⁰ Respiratory Medicine, Centre Hospitalier Universitaire de Liège—CHU, 4000 Liège, Belgium; j.guiot@chuliege.be
- * Correspondence: bart.vanrumste@kuleuven.be



Citation: Youssef Ali Amer, A.; Wouters, F.; Vranken, J.; Dreesen, P.; de Korte-de Boer, D.; van Rosmalen, F.; van Bussel, B.C.T.; Smit-Fun, V.; Duflot, P.; Guiot, J.; et al. Vital Signs Prediction for COVID-19 Patients in ICU. *Sensors* **2021**, *21*, 8131. <https://doi.org/10.3390/s21238131>

Academic Editor: Mario Munoz-Organero

Received: 8 November 2021

Accepted: 30 November 2021

Published: 5 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: This study introduces machine learning predictive models to predict the future values of the monitored vital signs of COVID-19 ICU patients. The main vital sign predictors include heart rate, respiration rate, and oxygen saturation. We investigated the performances of the developed predictive models by considering different approaches. The first predictive model was developed by considering the following vital signs: heart rate, blood pressure (systolic, diastolic and mean arterial, pulse pressure), respiration rate, and oxygen saturation. Similar to the first approach, the second model was developed using the same vital signs, but it was trained and tested based on a leave-one-subject-out approach. The third predictive model was developed by considering three vital signs: heart rate (HR), respiration rate (RR), and oxygen saturation (SpO₂). The fourth model was a leave-one-subject-out model for the three vital signs. Finally, the fifth predictive model was developed based on the same three vital signs, but with a five-minute observation rate, in contrast with the aforementioned four models, where the observation rate was hourly to bi-hourly. For the five models, the predicted measurements were those of the three upcoming observations (on average, three hours ahead). Based on the obtained results, we observed that by limiting the number of vital sign predictors (i.e., three vital signs), the prediction performance was still acceptable, with the average mean absolute percentage error (MAPE) being 12%, 5%, and 21.4% for heart rate, oxygen saturation, and respiration rate, respectively. Moreover, increasing the observation rate could enhance the prediction performance to be, on average, 8%, 4.8%, and 17.8% for heart rate, oxygen saturation, and respiration rate, respectively. It is envisioned that such models could be integrated with monitoring systems that could, using a limited number of vital signs, predict the health conditions of COVID-19 ICU patients in real-time.

Keywords: COVID-19; ICU; vital signs prediction; kNN-LS-SVM

1. Introduction

The recent COVID-19 pandemic shocked healthcare systems around the globe, highlighting the need for intelligent monitoring solutions. Intelligent monitoring solutions are needed to optimise the available resources in hospitals, more specifically in intensive care units (ICUs). During the first wave of the pandemic, ICUs experienced a shortage of beds, and healthcare workers were overloaded due to high admission rates and severely ill patients. Therefore, it became necessary to develop intelligent systems to optimise the available resources at hospitals, as well as the efforts of medical staff, to provide faster detection of patient deterioration. One key component of such an intelligent system involves data analytics of (big) medical data. Machine learning has become a popular and reliable analytical technique in recent years, especially in the medical domain. Many studies investigated hospitalised patients and ICU patients, for monitoring or mortality prediction [1–8]. Some of these studies investigated to what extent vital signs could inform on a patient's clinical deterioration and adverse events [1,2]. Other studies focused on the mortality prediction of ICU patients and the relevant predictors [3,4,6]. For instance, in our previous study [4], we investigated the mortality prediction of ICU patients based on a set of vital signs, namely, heart rate, blood pressure, respiration rate, and oxygen saturation. The observation rates of these vital signs were hourly to bi-hourly measurements for each one. The main concept in that study was to engineer interpretable/simple features in combination with a linear hard margin support vector machine that can inform mortality prediction for ICU patients one to two days in advance. Moreover, in the study by Mahdavi et al. [6], they developed support vector machine models to predict the mortality risk of COVID-19 patients, based on demographic and laboratory variables/features obtained from a patient's first day of admission. Another related approach involves vital sign predictors for hospitalised patients [7,8]. In our previous study [8], we developed predictive models to predict the future values of the same vital signs, up to three hours ahead on average. However, the population of that study comprised patients admitted to general wards (i.e., cardiology, post-surgical, and dialysis patients). Furthermore, in that study, the observations were sampled at 1 Hz using wearable devices (Somnotouch NIBP, <http://www.somnomedics.eu> (accessed on 28 November 2021)). Nevertheless, the acceptable predictions of both studies motivated us to integrate the two concepts, namely extracting simple features from low rate measurements of COVID-19 ICU patients to predict the near future values of the vital signs, up to three hours ahead on average.

In this study, we introduce, for the first time, machine learning predictive models to predict vital signs in COVID-19 patients. These machine learning models aim to predict a set of monitored vital signs, for patients in the ICU, in upcoming hours (i.e., three hours). Many machine learning algorithms are available that can be applied to such a case study. In their study, Han et al. [9] compared different machine learning algorithms (i.e., random forest, neural networks, and support vector machines) for a specific case study, and random forest outperformed. However, the machine learning algorithm we used in this study was the k-nearest neighbour least-squares support-vector machine (kNN-LS-SVM). This algorithm was chosen based on its efficient performance that was obtained in our previous study on vital signs predictors for hospitalised patients [8]. In addition, kNN-LS-SVM provides the advantage of localised learning algorithms, such as handling class imbalance, ambiguity, streaming analytics, and model personalisation [10]. The main focus of this study was to predict the future values of specific vital signs for upcoming observations (on average, three hours ahead) in COVID-19 ICU patients. These vital signs are mainly heart rate, respiration rate, and oxygen saturation. These vital signs were specified based on medical expert recommendations and constraints regarding the collected data. On the other hand, as portions of the collected data contain consistent measurements of blood pressure vital signs (i.e., systolic (SBP), diastolic (DBP), and mean arterial blood pressure (MAP)), we developed a predictive model considering these vital signs, as well for comparative reasons, to assess the influence of blood pressure (as a vital sign) on prediction performance.

2. Material and Methods

2.1. Data

Data used for developing and evaluating the models were collected for the WearIT4COVID Interreg project at the hospital Ziekenhuis Oost-Limburg (ZOL, Genk, Belgium), Maastricht University hospital MUMC+ (Maastricht, The Netherlands), and Le Centre hospitalier universitaire de Liège (CHU) (Liège, Belgium), during the first and second waves of the COVID-19 pandemic. The data were collected from the three hospitals with the details shown in Table 1. For ZOL, the data of COVID-19 positive patients admitted at the intensive care unit (ICU) were collected based on approval of the local ethics committee (20/003R). A total of 51 patients were included in this dataset with various lengths of stay at the ICU. The mean duration of ICU stay was 10.6 days (± 8.87 days), and the mean of the entire hospital stay was 16.7 days (± 9.27). For MUMC+, data collection in the Maastricht Intensive Care COVID (MaastrICChT) cohort was carried out by researchers from the ICU department. Ethical approval was obtained from the Medical Ethical Committee UM/azM (Maastricht). All patients diagnosed with COVID-19, and entering the ICU department, were included in the study; data were collected until patients were discharge from the ICU. Data from 50 patients were provided for analysis. The mean duration of ICU stay was 16.9 days (± 16.3). For CHU Liège, the data of 84 patients were collected, and patients who were monitored for less than 24 h were excluded. The collected data from the three hospitals contained the vital parameters, demographic data (such as age and gender), comorbidities, and outcomes of 185 patients in total.

Table 1. The demographic presentation and most frequent comorbidities (prior to or acquired during their COVID infection) of the patients admitted to ICU in the three hospitals: ZOL, MUMC+, and CHU de Liège.

	ZOL (51 Patients)	MUMC+ (50 Patients)	CHU de Liège (84 Patients)
Demographic and comorbidity parameters	Descriptive statistics		
Age (mean \pm std)	65.84 \pm 12.44	67.02 \pm 1.28	69 \pm 12.16
Gender (male, %)	62.7%	80.0%	68%
Height (cm; mean \pm std)	166.63 \pm 12.11	176.3 \pm 8.3	168 \pm 26.53
Weight (kg; mean \pm std)	83.76 \pm 16.34	85.91 \pm 13.68	81.45 \pm 20.01
Smoking status (%)	Never: 66% Smoker: 6.4% Former smoker: 27.7%	Never: 90% Smoker: 6.0% Former smoker: 4.0%	Never: 50% Smoker: 6% Former Smoker: 12%
Cardiovascular disease (%)	17.6%	4.0 %	23% (yes) 22% (no) 55% (unknown)
Diabetes (%)	27.5%	18.0 %	51% (yes) 21% (no) 28% (unknown)
Arterial hypertension (%)	43.1%	16.0%	61% (yes) 14% (no) 25% (unknown)
Cerebrovascular accident or transient ischaemic attack (%)	5.9%	2.0%	(unknown)
Kidney insufficiency (%)	60.8%	0.0%	(unknown)
Heart Failure (%)	29.4%	4.0%	(unknown)
NYHA classification	II: 8.3% III: 8.3%	6.0%	(unknown)
Myocardial infarction (%)	9.8%	42%	(unknown)
PCI/PTCA	9.8%	67.02 \pm 1.28	(unknown)
COPD	11.8%	80.0%	24% (yes) 29% (no) 47% (unknown)
Asthma (%)	7.8%	176.3 \pm 8.3	9% (yes) 38% (no) 53% (unknown)
Intubated	23 (46%)	50 (100%)	38 (46%)
O ₂ Mask/Nasal Cannula	40 (78%)	(Unknown)	71 (86%)

2.2. Methods

In this section, we address the machine learning algorithm that was used to develop the predictive model. The used machine learning algorithm was the hybrid localised learning algorithm k nearest neighbour least-squares support-vector machine (kNN-LS-SVM) for regression. This algorithm was introduced in our previous article [8] to predict the future values of the monitored vital signs of hospitalised patients in general wards using wearable sensors. The kNN-LS-SVM algorithm for regression was developed based on the studies of [10–17]. The integration of the kNN algorithm into the LS-SVM algorithm was proposed to build an efficient predictive model (LS-SVM) within a local region, which is determined by the kNN search algorithm. Localising the model is considered beneficial, to consider the local characteristics of the investigated region. In addition, such a localised learning approach could help to build updated models and provide online analytics [10].

Local Learning of SVMs

In this section, we begin by reviewing the main concepts behind support vector machines (SVMs) and localised learning approaches for SVMs.

SVMs are originally presented as binary classifiers that assign each datum instance $\mathbf{x} \in \mathbb{R}^d$ to one of two classes described by a class label $y \in \{-1, 1\}$, based on the decision boundary that maximises the margin $2/\|\mathbf{w}\|_2$ between the two classes. Generally, a feature map $\phi: \mathbb{R}^d \mapsto \mathbb{R}^p$ is used to transform the geometric boundary between the two classes to a linear boundary $L: \mathbf{w}^\top \phi(\mathbf{x}) + b = 0$ in a feature space, for some weight vectors $\mathbf{w} \in \mathbb{R}^{p \times 1}$ and $b \in \mathbb{R}$. The class of each instance can then be found by $y = \text{sgn}(\mathbf{w}^\top \phi(\mathbf{x}) + b)$, where sgn refers to the sign function.

Similar to the classification problems, regression models are obtained via estimating the boundary L based on a set of training examples \mathbf{x}_i ($1 \leq i \leq N$) with corresponding output values $y_i \in \mathbb{R}$. In particular, one is interested in parameters \mathbf{w} and b that minimise a loss-function:

$$\min_{\mathbf{w}, b; \zeta} \frac{1}{2} \mathbf{w}^\top \mathbf{w} + C \sum_{i=1}^N (\zeta_i + \zeta_i^*), \quad (1)$$

and are subject to:

$$\begin{aligned} y_i - \mathbf{w}^\top \phi(\mathbf{x}_i) - b &\leq \epsilon + \zeta_i, & i = 1, 2, \dots, N, \\ \mathbf{w}^\top \phi(\mathbf{x}_i) - y_i + b &\leq \epsilon + \zeta_i^*, & i = 1, 2, \dots, N, \\ \zeta_i, \zeta_i^* &\geq 0, & i = 1, 2, \dots, N. \end{aligned}$$

The constant C in (1) denotes the *penalty term* that is used to penalise estimation error through the slack variables ζ_i and ζ_i^* outside Vapnik ϵ -sensitivity loss function in the optimisation process.

LS-SVM's are obtained by using a least-squares error loss function [16]:

$$\min_{\mathbf{w}, b; e} \frac{1}{2} \mathbf{w}^\top \mathbf{w} + \frac{1}{2} \gamma \sum_{i=1}^N e_i^2, \quad (2)$$

such that

$$y_i = \mathbf{w}^\top \phi(\mathbf{x}_i) + b + e_i, \quad i = 1, 2, \dots, N.$$

where γ is the regularisation constant for LS-SVM. The optimisation procedure introduces errors e_i , such that $1 - e_i$ is proportional to the signed distance of \mathbf{x}_i from the decision boundary. In fact, the non-negative slack variable constraint is removed and the solution of the optimisation problem can be obtained by a set of linear equations, reducing computational efforts [16].

Local learning approaches build models that fit the data in the local neighbourhood around a test example and by locally adjusting the model parameters to the properties of the data [14].

While global SVMs consider the same weights for all training instances in the optimisation process (2), local learning approaches allow for the training samples near a test point to be more influential than others. Localised learning approaches of SVMs [10] are based on weighting functions $\lambda(\mathbf{x}_s, \mathbf{x}_i)$ that express the similarity between the feature vectors of the i -th data point \mathbf{x}_i and a test instance \mathbf{x}_s . For an LS-SVM, this leads to the following cost function:

$$\min_{\mathbf{w}, b, e} \frac{1}{2} \mathbf{w}^\top \mathbf{w} + \frac{1}{2} \gamma \sum_{i=1}^N \lambda(\mathbf{x}_s, \mathbf{x}_i) e_i^2, \quad (3)$$

such that

$$y_i = \mathbf{w}^\top \boldsymbol{\phi}(\mathbf{x}_i) + b + e_i, \quad i = 1, 2, \dots, N.$$

where y_i is a real number. Weighted least-squares support vector machines [17] use a similar approach, but here, a different weighting function can be used for any given test point \mathbf{x}_s . In this work, we study a binary valued similarity criterion:

$$\lambda(\mathbf{x}_s, \mathbf{x}_i) = \begin{cases} 1 & \text{if } \|\mathbf{x}_s - \mathbf{x}_i\|_2 \leq r_s \\ 0 & \text{otherwise,} \end{cases}$$

where r_s is the K -th smallest distance among $\{\|\mathbf{x}_s - \mathbf{x}_i\|_2; 1 \leq i \leq N\}$. This formulation leads to the hybrid KNN-LS-SVM method that we applied on the time-series prediction approach. In particular, a regression model was built for each test example using only the training examples located in the vicinity of the test example [11].

KNN-LS-SVM has the additional advantage of sparseness. Indeed, for an LS-SVM or the localised version that uses a continuous similarity function, all input data are required to construct the separating hyperplane [17]. This can be seen by solving the optimisation problem (3). Using the method of the Lagrangian multipliers, we find:

$$\mathcal{L}(w, b, e; \alpha) = \frac{1}{2} \|w\|_2^2 + \frac{1}{2} \gamma \sum_{i=1}^N \lambda(\mathbf{x}_s, \mathbf{x}_i) e_i^2 - \sum_{i=1}^N \alpha_i (\mathbf{w}^\top \boldsymbol{\phi}(\mathbf{x}_i) + b + e_i - y_i),$$

where α_i are the *Lagrangian* multipliers. Thus, for a KNN-LS-SVM, the sparseness characteristic is returned to the LS-SVM, as only the neighbour data points contribute to the model. In an online learning mode, this sparseness will result in a computational advantage compared to LS-SVM.

The illustration of the kNN-LS-SVM regression algorithm is shown in Figure 1. The algorithm of KNN-LS-SVM is implemented as follows:

1. Given a test example \mathbf{x}_s , compute distances to all training examples and pick the nearest K neighbours.
2. Train the LS-SVM model with the K nearest neighbours.
3. Use the resulting regressor to estimate the output of \mathbf{x}_s .

The parameter K and the distance metric (e.g., Euclidean, Mahalanobis, or Chebyshev) are additional hyperparameters next to the kernel width σ_0 and the penalty term γ that are optimised in a cross-validation approach [10]. One challenge faced in finding the nearest neighbour in a continuously increasing data pool is the search complexity. However, several advanced search algorithms were developed to reduce this complexity [18].

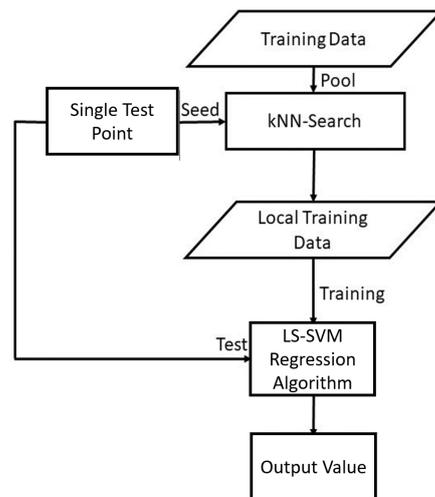


Figure 1. A flow chart illustrating the localised learning algorithm of KNN-LS-SVM for regression (adapted from ref. [8]).

3. Results

In this section, we introduce the prediction results of the developed predictive models following two test approaches. The first approach is the 80/20 approach, in which 80% (7100 datapoints) of the shuffled data are used to train the model, and the remaining 20% (1845 datapoints) of the data are used to test the model. The other approach is leave-one-subject-out, in which for each subject, this subject's data (on average 50 datapoints) is excluded from the training set of the model to be used as the test set.

The developed models in this study are based on data from the three hospitals. Moreover, the results from the three hospitals together are consistent with the results from each hospital individually. Therefore, in this section, we focus on the results from the merged data of the three hospitals together. Prior to applying the localised learning approach of kNN-LS-SVM, a set of statistical features were extracted from the raw measurements of each vital sign. The extracted features were the minimum, maximum, mean, median, standard deviation, and energy. These features were extracted from a time window of three observations to predict the upcoming vital sign values up to three observations ahead (approximately three hours on average, as the observation rate is not uniform). After extracting the aforementioned features, we developed five predictive models. The predictive models were evaluated based on the mean absolute percentage error (MAPE) which is calculated by the formula $MAPE = \frac{1}{n} \sum_{t=1}^n \|(A_t - P_t)/A_t\|$, where n is the number of observations, A_t and P_t are the actual and predicted values at time t .

For the first predictive model, we aimed to predict the future values of the vital signs for the upcoming three observations (on average three hours ahead) based on hourly to bi-hourly observation rate. These vital signs were heart rate, respiration rate, oxygen saturation, and blood pressure (systolic, diastolic, and mean arterial). For this predictive model, the features were extracted from all of these vital signs and forming vectors, with (6 features \times 6 vital signs) 36 dimensions. In other words, all vital signs contributed to predicting each vital sign. The prediction performance is shown in Figure 2 in terms of mean absolute percentage error (MAPE).

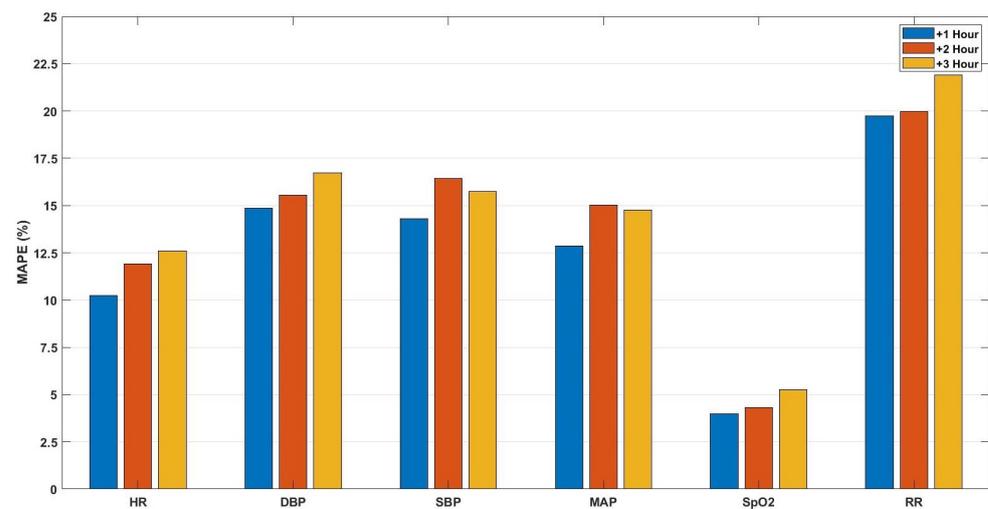


Figure 2. The mean absolute percentage error of the predictions for the upcoming three hours (+1, +2, +3 h) of the vital signs: heart rate (HR), diastolic (DBP), systolic (SBP), mean arterial blood pressure (MAP), oxygen saturation (SpO₂), and respiration rate (RR), respectively, based on the hourly to bi-hourly observation rate.

The second predictive model is similar to the first, however, it is trained/tested based on the leave-one-subject-out approach. As shown in Figure 3, the box-plots of the resulting MAPE for the predictions of the six vital signs are depicted.

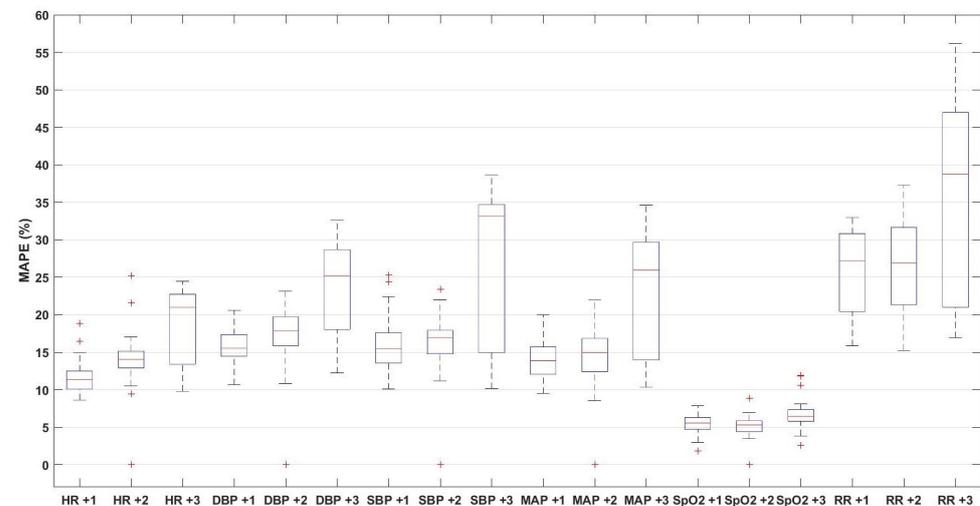


Figure 3. The box-plot of leave-one-subject-out predictions of the vital signs of heart rate (HR), diastolic (DBP), systolic (SBP), mean arterial blood pressure (MAP), oxygen saturation (SpO₂), and respiration rate (RR) for the upcoming three observations (on average 3 h ahead).

For the third predictive model, we aimed to predict the future values of only three vital signs for the upcoming three observations (three hours on average). These vital signs were heart rate, respiration rate, oxygen saturation. These three vital signs were selected as they were recorded in a more consistent way compared to blood pressure. Furthermore, these vital signs could be easily measured using wearable sensors in contrast with blood pressure. For this predictive model, the features were extracted from the three vital signs as well forming vectors of (6 features \times 3 vital signs) 18 dimensions. The prediction performance is shown in Figure 4 in terms of MAPE. Moreover, examples of the comparison between the predicted and the actual values of the three vital signs for the upcoming three hours are shown in Figure 5.

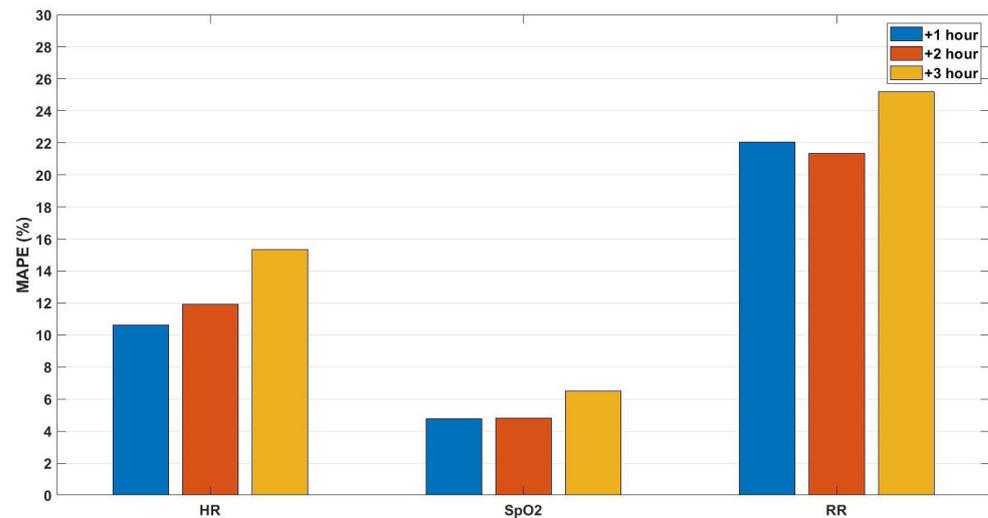


Figure 4. The mean absolute percentage error of the predictions for the upcoming three hours (+1, +2, +3 h) of the vital signs: heart rate (HR), oxygen saturation (SpO₂), and respiration rate (RR), respectively, based on hourly to bi-hourly observation rate.

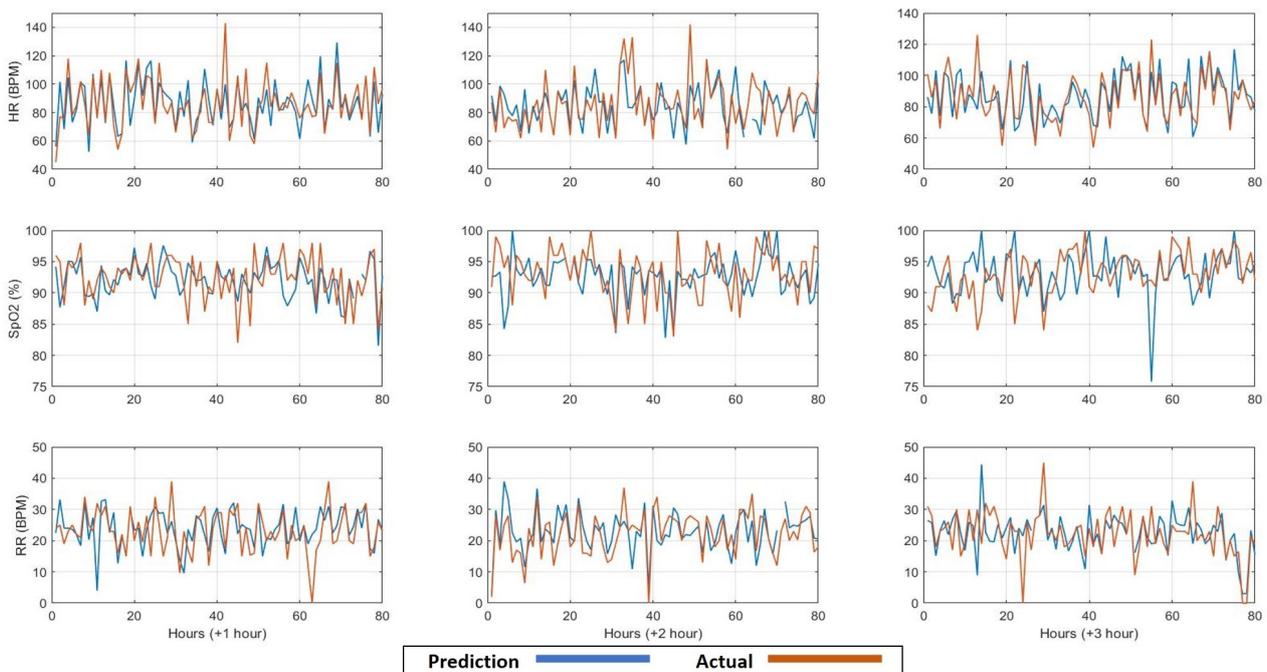


Figure 5. The comparison between the predicted and the actual values for the upcoming three hours (+1, +2, +3 h) of the vital signs: heart rate (HR), oxygen saturation (SpO₂) and respiration rate (RR), respectively, based on hourly to bi-hourly observation rate.

Similar to the third model, for the fourth predictive model, it was based on the three vital signs (heart rate, respiration rate, and oxygen saturation), but it was trained and tested using the leave-one-subject-out approach. The box-plot of the resulting MAPE of the predictions are shown in Figure 6.

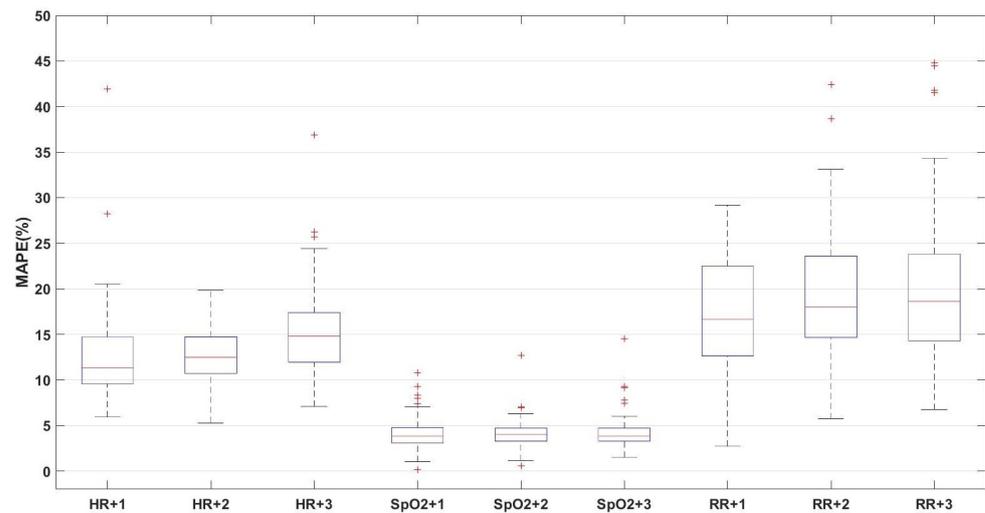


Figure 6. The box-plot of the leave-one-subject-out prediction of the vital signs of heart rate (HR), oxygen saturation (SpO₂), and respiration rate (RR) for the upcoming three observations (on average 3 h ahead).

Finally, the fifth predictive model predicted the future values of heart rate, oxygen saturation, and respiration rate for the upcoming three observations (on average three hours). However, the observation rate, in this case, was one observation/5 min. Therefore, the same features were extracted from the same time windows, but with a larger number of observations (i.e., 12 observations per hour), forming vectors with $(3 \text{ h} \times 6 \text{ features} \times 3 \text{ vital signs})$ 54 dimensions to predict the 1-hour average values for the upcoming three hours. The prediction performance is shown in Figure 7 in terms of MAPE. Furthermore, in Figure 8, we show some examples of the comparison between the predicted and the actual values of the three vital signs with observation rate, one observation/5 min, for the upcoming three hours.

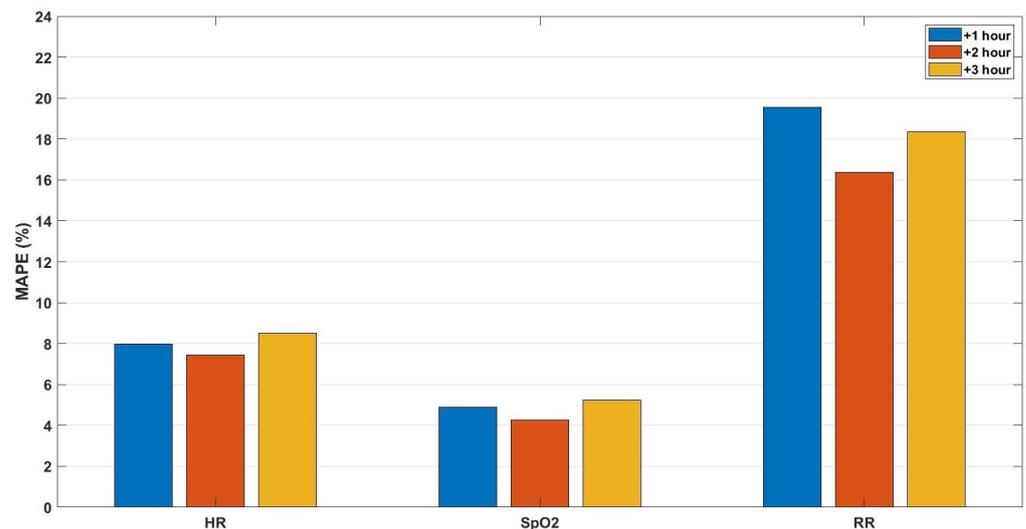


Figure 7. The mean absolute percentage error of the predictions for the upcoming three hours (+1, +2, +3 h) of the vital signs: heart rate (HR), oxygen saturation (SpO₂), and respiration rate (RR), respectively, based on an observation rate of one observation/5 min.

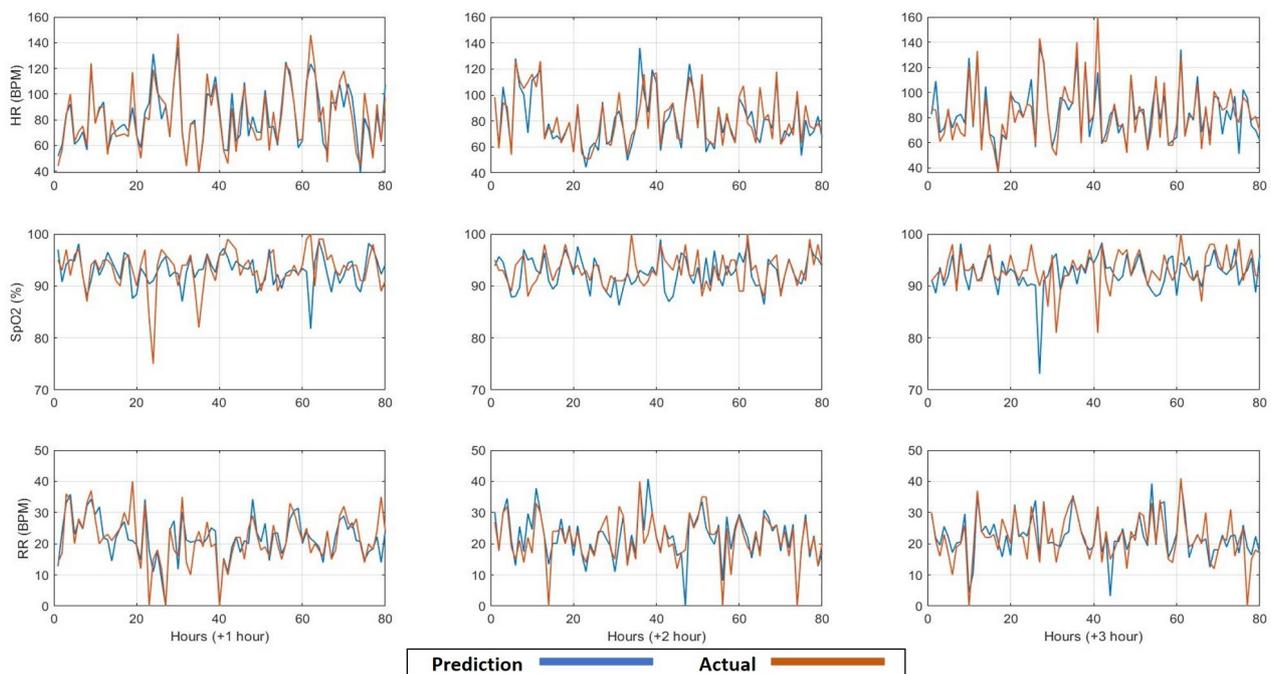


Figure 8. The comparison between the predicted and the actual values for the upcoming three hours (+1, +2, +3 h) of the vital signs: heart rate (HR), oxygen saturation (SpO₂), and respiration rate (RR), respectively, based on observation rate of one observation/5 min.

Based on the obtained predictions, the histograms of the obtained error between the predicted and actual values were calculated for both the third and fifth predictive models (cf. supra) as depicted in Figures 9 and 10 for the three vital signs (heart rate, oxygen saturation, and respiration rate) for the upcoming three hours. After performing the prediction, the early warning score components for the predicted vital signs were calculated for the predicted and the actual values. The absolute error between the predicted and the actual EWS components was calculated for the third and fourth predictive models based on the ranges depicted in Table 2. The normalised histograms of the absolute EWS error are shown in Figures 11 and 12 for the third and fifth predictive models, respectively.

Table 2. Early warning score system based on Ziekenhuis Oost-Limburg (ZOL) Hospital.

SCORE	3	2	1	0	1	2	3
Temperature (°C)		<35.1	35.1–36.5	36.6–37.5	>37.5		
Heart Rate (BPM)		<40	40–50	51–100	101–110	111–130	>130
Respiration Rate (BPM)		<9		9–14	15–20	21–30	>30
Oxygen Saturation (%)	<91	91–93	94–95	>95			
Systolic Blood Pressure (mmHg)	<70	70–80	81–100	101–180	180–200	>200	

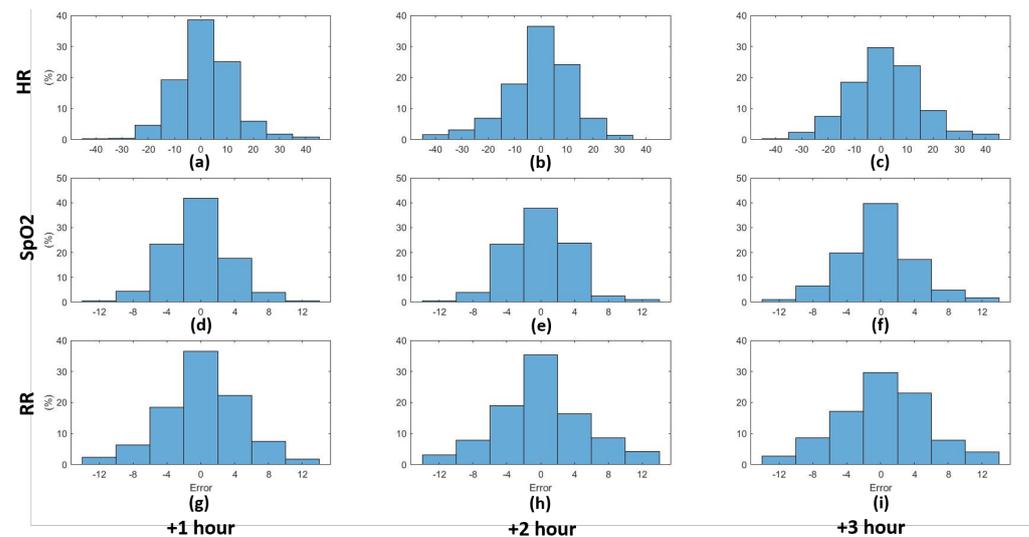


Figure 9. Normalised histograms of the error of the predicted values for the three vital signs for the upcoming three hours based on hourly to bi-hourly observations. Each row represents one of the three vital signs: heart rate, oxygen saturation, and respiration rate respectively. Each column represents the prediction hour: +1, +2, +3 h, respectively. (a) HR +1 h, (b) HR +2 h, (c) HR +3 h, (d) SpO₂ +1 h, (e) SpO₂ +2 h, (f) SpO₂ +3 h, (g) RR +1 h, (h) RR +2 h, (i) RR +3 h.

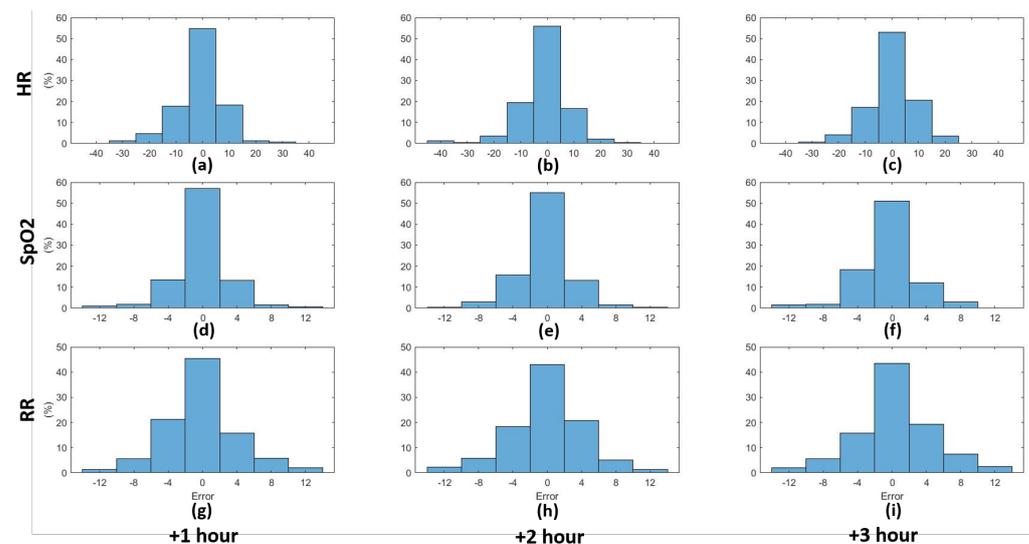


Figure 10. Normalised histograms of the error of the predicted values for the three vital signs for the upcoming three hours based on observation rate of one observation/5 min. Each row represents one of the three vital signs: heart rate, oxygen saturation, and respiration rate respectively. Each column represents the prediction hour: +1, +2, +3 h, respectively. (a) HR +1 h, (b) HR +2 h, (c) HR +3 h, (d) SpO₂ +1 h, (e) SpO₂ +2 h, (f) SpO₂ +3 h, (g) RR +1 h, (h) RR +2 h, (i) RR +3 h.

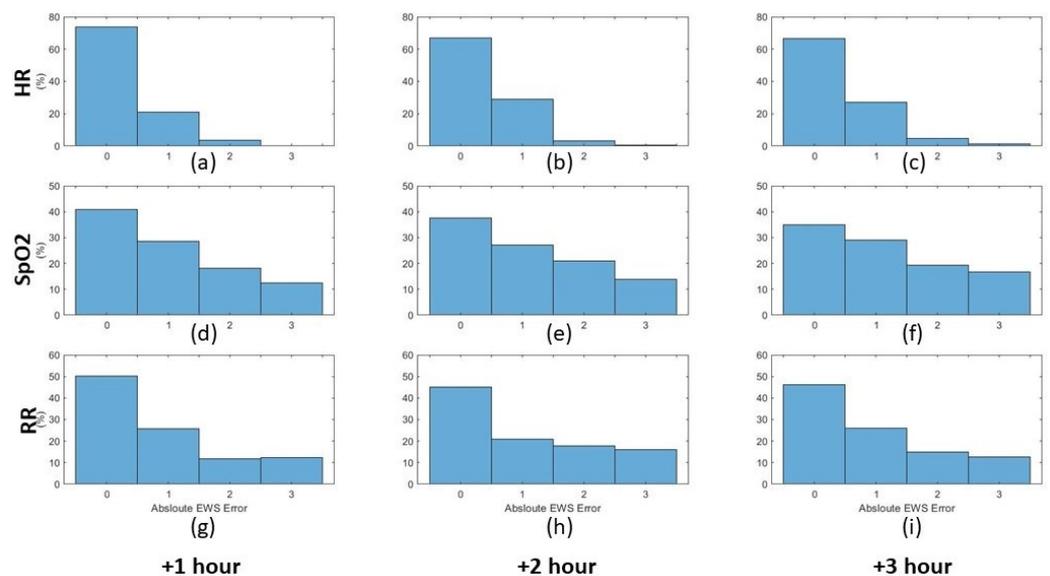


Figure 11. Normalised histograms of the absolute EWS error of the predicted values for the three vital signs for the upcoming three hours based on hourly to bi-hourly observations. Each row represents one of the three vital signs: heart rate, oxygen saturation, and respiration rate, respectively. Each column represents the prediction hour: +1, +2, +3 h, respectively. (a) HR +1 h, (b) HR +2 h, (c) HR +3 h, (d) SpO₂ +1 h, (e) SpO₂ +2 h, (f) SpO₂ +3 h, (g) RR +1 h, (h) RR +2 h, (i) RR +3 h.

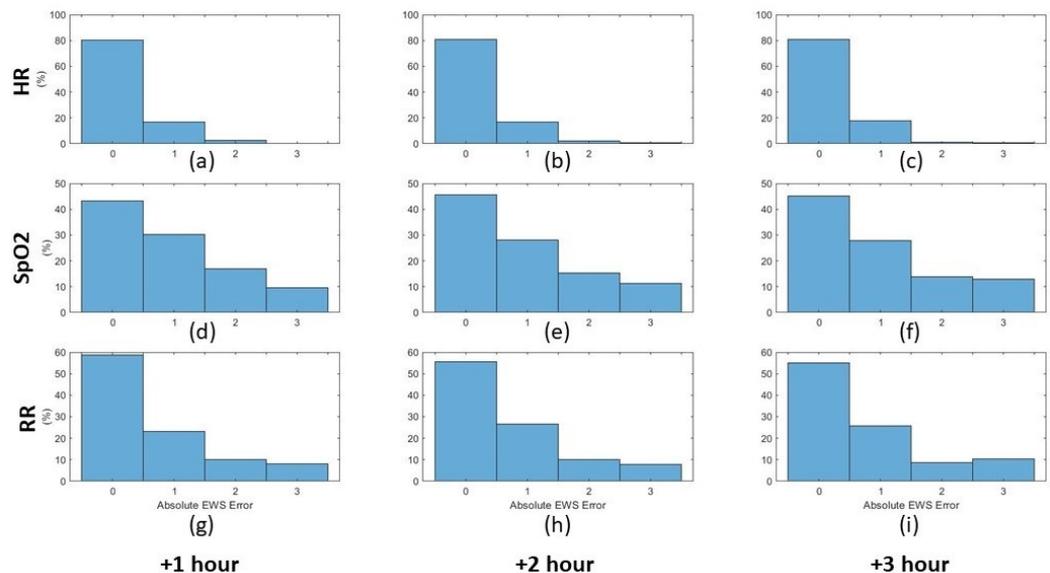


Figure 12. Normalised histograms of the absolute EWS error of the predicted values for the three vital signs for the upcoming three hours based on the observation rate of one observation/5 min. Each row represents one of the three vital signs: heart rate, oxygen saturation, and respiration rate, respectively. Each column represents the prediction hour: +1, +2, +3 h, respectively. (a) HR +1 h, (b) HR +2 h, (c) HR +3 h, (d) SpO₂ +1 h, (e) SpO₂ +2 h, (f) SpO₂ +3 h, (g) RR +1 h, (h) RR +2 h, (i) RR +3 h.

4. Discussion

Based on the obtained results, we can compare the performance of the different models. For example, by comparing the first and second predictive models, it is noticeable that the performance of the first model, in terms of MAPE, is better for all vital signs than the median MAPE values of the second model, which indicates that the performance is better when the same-person data are considered in the training set.

By comparing the prediction performance of the first and third models, as both are based on hourly to bi-hourly observation rates, we notice that the HR and SpO₂ prediction error for the first model is similar to that of the third model for the first and second-hour predictions, without a statistical significance (at 5% significance level), and less for the third hour without significance. This similarity can indicate that, for these models, the considered blood pressure vital signs are unnecessary for both HR and SpO₂. However, for RR, the prediction performance of the first model (MAPE = 19.8%, 20%, and 22%) is better than that of the third model (MAPE = 22%, 21.3%, 25.2%) with significance at 10% significance level. This outperformance indicates that blood pressure vital signs can be informative to predict the respiration rate. Given these results, we can conclude that blood pressure vital signs did not significantly influence the prediction performance for both heart rate and oxygen saturation in this setup. On the other hand, respiration rate prediction is influenced by the absence of blood pressure parameters. However, we should notice that the respiration rate for the considered population is biased because at some moment(s) during their stays, patients were coupled to mechanical ventilation, which makes the prediction of respiration rate rather ambiguous. Moreover, it is worth mentioning that for all predictive models, the MAPE values of RR are with the highest values, as the range of RR is small compared to other vital signs (15 ± 6 BPM). Hence, an error with 1 BPM represents approximately 5% of the original value.

For the third and fourth models, we observed that the performance of both models is approximately similar without statistical significance by comparing the average MAPE values in Figure 4 and the average values of the box plots in Figure 5. This result indicates that the models based on the three vital signs are not notably influenced by the same-person data.

By comparing the performance of the third and fifth models, we observe a significant enhancement in the prediction performance of the fifth model compared to the third model, as shown in Figures 4 and 6, which indicates that by increasing the observation rate, the prediction performance will enhance.

A general note for SpO₂ prediction is that the variation range of SpO₂ is between 85 to 100% for this population. Therefore, an absolute percentage error of 5% represents almost one-third of the variation range, which can be considered as a limitation of the prediction performance.

Furthermore, the normalised histograms of the prediction errors as shown in Figures 9 and 10 show the outperformance of the fifth model compared to the third model. More specifically, for HR, the error frequency within the lowest error range (−5 to 5 BPM) is approximately 20% higher for the fifth model than for the third model, which indicates a lower error frequency for the higher error values. Similarly, for both SpO₂ and RR, the error frequency within the lowest error range (−2 to 2%) and (−2 to 2 BPM) respectively, is higher for the fifth model than the third model with approximately 10%. Furthermore, by applying the F-test to test the significance between the different distribution variances, it is observed that the significance (5% significance level) is observed between each distribution in Figure 9 and its equivalent in Figure 10. This significance approves the improvement in the error distributions that can be obtained by increasing the observation rate. Moreover, by comparing the absolute EWS error between the two models, it is noticeable that the prediction error for the third model is lower than that of the second model for the three parameters for the three prediction horizons, as shown in Figures 11 and 12.

5. Conclusions

Conclusively, in this study, we introduced for the first time machine learning predictive models to predict the future values of COVID-19 patients' vital signs at ICU. These predictive models have shown promising performances with a few parameters (i.e., three vital signs) in predicting their values for the upcoming three hours. This performance is evaluated in terms of MAPE and it is on average 12%, 5% and 21.4% for HR, SpO₂, and RR respectively. Moreover, the performances of the predictive models based on three vital

signs show robustness in training the model, whether considering the same-person data or not. Ultimately, increasing the observation rate could enhance the prediction performance to be on average 8%, 4.8%, and 17.8% for heart rate, oxygen saturation, and respiration rate, respectively. Ultimately, the obtained results show promising performance for the particular profile of COVID-19 ICU patients.

For future work, having a high data rate (e.g., 1 Hz) may enhance the performance. Moreover, we would suggest to test our approach on COVID-19 patients outside ICU for real-time monitoring. Such real-time monitoring and prediction can lighten the workload on the medical staff in critical times during the pandemic. Furthermore, for larger datasets, dividing the group of patients into sub-groups based on their comorbidities may enhance the prediction performance. Ultimately, one of the main hurdles that may make it challenging to provide real-time analytics is the limited data storage resources at most hospitals. Hence, upgrading data storage facilities at hospitals would be a game-changer for intelligent medical systems.

Author Contributions: Conceptualisation, A.Y.A.A., B.V., and J.-M.A.; Data curation, J.V., F.W., D.d.K.-d.B., V.S.-F., P.D. (Patrick Dufлот), P.V., F.v.R., B.C.T.v.B. J.G., I.C.C.v.d.H. and D.M.; Formal analysis, A.Y.A.A.; Investigation, A.Y.A.A.; Methodology, A.Y.A.A.; Resources, B.V., J.-M.A.; Software, A.Y.A.A.; Supervision, B.V., and J.-M.A.; Validation, A.Y.A.A.; Visualisation, A.Y.A.A.; Drafting of protocol EAGLE study, J.V., D.d.K.-d.B., V.S.-F., P.D. (Patrick Dufлот); Writing—original draft, A.Y.A.A.; Writing—review and editing, A.Y.A.A., J.V., F.W., D.d.K.-d.B., F.v.R., B.C.T.v.B., P.D. (Pauline Dreesen), P.V., B.V., and J.-M.A. All authors have read and agreed to the published version of the manuscript.

Funding: The WearIT4COVID project, carried out within the framework of the Interreg V-A Euregio Meuse-Rhine programme, is financed (EUR 1.1 million) by the European Union and the European Regional Development Fund (ERDF), as well as project partners. By investing European funds in Interreg projects, the European Union is investing directly in economic development, innovation, territorial development, social inclusion, and education in the Euregio Meuse-Rhine.

Institutional Review Board Statement: This study was approved by the Medical Ethical Research Assessment Committee (METC) of Maastricht UMC+ (approval number 2020-1565/300523) and registered in the Netherlands Trial Register (NL8613). It was conducted in accordance with the Declaration of Helsinki. During the pandemic, the board of directors of Maastricht UMC+ adopted a policy to inform patients and ask their consent to use the collected data for COVID-19 research purposes, and to register consent in the healthcare information system instead of signing a paper consent form. This study was approved by the Clinical Trial Unit and Ethical Committee of Ziekenhuis Oost-Limburg (approval number 20/0030R). It was conducted in accordance with the Declaration of Helsinki, the ICH-GCP standards and GDPR regulation.

Informed Consent Statement: Patient consent was waived due to provision in GDPR regulation article 9.4 and 89 translated into Belgian law of 30 July 2018 relating to the protection of individuals with regard to the processing of personal data. Pseudonimisation was used as a data minimisation mean. No informed consent was obtained as this study is not subjected to the law of the 7th of May 2004 ‘Law concerning experiments on humans’.

Data Availability Statement: Dataset supporting the study may be obtained upon request to Liege University Hospital, Maastricht UMC+, and Ziekenhuis Oost-Limburg in accordance with applicable laws.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Brekke, I.J.; Puntervoll, L.H.; Pedersen, P.B.; Kellett, J.; Brabr, M. The value of vital sign trends in predicting and monitoring clinical deterioration: A systematic review. *PLoS ONE* **2019**, *14*, e0210875. [[CrossRef](#)] [[PubMed](#)]
2. Kause, J.; Smith, G.; Prytherch, D.; Parr, M.; Flabouris, A.; Hillman, K. A comparison of Antecedents to Cardiac Arrests, Deaths and Emergency Intensive care Admissions in Australia and New Zealand, and the United Kingdom—The ACADEMIA study. *Resuscitation* **2004**, *62*, 275–82. [[CrossRef](#)] [[PubMed](#)]

3. Barfod, C.; Lauritzen, M.M.P.; Danker, J.K.; Sölétormos, G.; Forberg, J.L.; Berlac, P.A.; Lippert, F.; Lundstrøm, L.H.; Antonsen, K.; Lange, K.H.W. Abnormal vital signs are strong predictors for intensive care unit admission and in-hospital mortality in adults triaged in the emergency department—A prospective cohort study. *Scand. Trauma Resusc. Emerg. Med.* **2012**, *10*, 20–28. [[CrossRef](#)] [[PubMed](#)]
4. Youssef Ali Amer, A.; Vranken, J.; Wouters, F.; Mesotten, D.; Vandervoort, P.; Storms, V.; Luca, S.; Vanrumste, B.; Aerts, J.-M. Feature engineering for ICU mortality prediction based on hourly to bi-hourly measurements. *Appl. Sci.* **2019**, *9*, 3525. [[CrossRef](#)]
5. Redfern Oliver C.; Pimentel, M.A.F.; David, P.; Meredith, P. Predicting in-hospital mortality and unanticipated admissions to the intensive care unit using routinely collected blood tests and vital signs: Development and validation of a multivariable model. *Resuscitation* **2018**, *133*, 75–81. [[CrossRef](#)] [[PubMed](#)]
6. Mahdavi, M.; Choubdar, H.; Zabehe, E.; Rieder, M.; Safavi-Naeini, S. A machine learning based exploration of COVID-19 mortality risk. *PLoS ONE* **2021**, *16*, e0252384. [[CrossRef](#)] [[PubMed](#)]
7. Liu, S.; Yao, J.; Motani, M. Early Prediction of Vital Signs Using Generative Boosting via LSTM Networks. In Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 18–21 November 2019.
8. Youssef Ali Amer, A.; Wouters, F.; Vranken, J.; de Korte-de Boer, D.; Smit-Fun, V.; Dufloot, P.; Beaupain, M.-H.; Vandervoort, P.; Luca, S.; Aerts, J.-M.; et al. Vital Signs Prediction and Early Warning Score Calculation Based on Continuous Monitoring of Hospitalised Patients Using Wearable Technology. *Sensors* **2020**, *20*, 6593. [[CrossRef](#)] [[PubMed](#)]
9. Han, T.; Jiang, D.; Zhao, Q.; Wang, L.; Yin, K. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Trans. Inst. Meas. Control* **2018**, *40*, 2681–2693. [[CrossRef](#)]
10. Youssef Ali Amer, A.; Aerts, J.; Vanrumste, B.; Luca, S. A Localised Learning Approach Applied to Human Activity Recognition. *IEEE Intell. Syst.* **2020**, *99*, 58–71.
11. Youssef Ali Amer, A. Localised Least Squares Support Vector Machines with Application to Weather Forecasting. Master's Thesis, KU Leuven, Leuven, Belgium, 2016.
12. Cheng, H.; Tan, P.-N.; Jin, R. Localized support vector machine and its efficient algorithm. In Proceedings of the SIAM International Conference on Data Mining, Minneapolis, MN, USA, 26–28 April 2007.
13. Cheng, H.; Tan, P.; Jin, R. Efficient algorithm for localized support vector machine. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 537–549. [[CrossRef](#)]
14. Bottou, L.; Vapnik, V. Local Learning Algorithms. *Neural Comput.* **1992**, *4*, 888–900. [[CrossRef](#)]
15. Yang, H.; Huang, K.; King, I.; Lyu, M.R. Localized support vector regression for time series prediction. *Neurocomputing* **2009**, *72*, 10–12. [[CrossRef](#)]
16. Suykens, J.A.K.; Vandewalle, J. Least Squares Support Vector Machine Classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [[CrossRef](#)]
17. Suykens, J.A.K.; De Brabanter, J.; Lukas, L.; Vandewalle, J. Weighted least squares support vector machines: Robustness and sparse approximation. *Neurocomputing* **2002**, *48*, 85–105. [[CrossRef](#)]
18. Cayton, L. Fast nearest neighbor retrieval for bregman divergences. In Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, 5–9 July 2008; pp. 112–119, ISBN 9781605582054. [[CrossRef](#)]