

Estimation de la pression partielle artérielle du dioxyde de carbone par des paramètres noninvasifs

par

Jihad EL TANNOURY

MÉMOIRE PAR ARTICLE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE
SUPÉRIEURE COMME EXIGENCE PARTIELLE À L'OBTENTION
DE LA MAÎTRISE AVEC MÉMOIRE EN GÉNIE
CONCENTRATION PERSONNALISÉE
M. Sc. A.

MONTRÉAL, LE 26 MARS 2020

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC



Jihad Tannoury, 2020



Cette licence Creative Commons signifie qu'il est permis de diffuser, d'imprimer ou de sauvegarder sur un autre support une partie ou la totalité de cette oeuvre à condition de mentionner l'auteur, que ces utilisations soient faites à des fins non commerciales et que le contenu de l'oeuvre n'ait pas été modifié.

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE:

Mme Rita Noumeir, directrice de mémoire
Département de génie électrique, École de technologie supérieure

M. Philippe Jovet, co-directeur
Centre de recherche CHU, Hopital Sainte Justine

M. Luc Duong, président du jury
Département de génie logiciel et des TI, École de technologie supérieure

M. Jean-Marc Lina, examinateur externe
Département de génie électrique, École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 3 MARS 2020

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je tiens à remercier avant tout mes parents, ma soeur Chantal qui a déjà parcouru ce chemin avec un grand succès et mon frère Ziad qui va y entrer bientôt, je les remercie pour leur support continu durant mon séjour à Montréal ; je suis sûr que sans leur implication, ce travail n'aurait pas pu se réaliser, je leur serais toujours reconnaissant.

Je tiens également à remercier ma directrice de recherche, professeure Rita Noumeir, pour accepter de m'encadrer durant mes travaux dans le laboratoire de traitement de l'information en santé LATIS. Ton soutien, ton encouragement et tes conseils judicieux ont facilité beaucoup mon travail. J'aimerais aussi remercier docteur Philippe Jovet, mon codirecteur, pour sa disponibilité, ses remarques et ses commentaires constructifs tout au long de ma maîtrise.

Il faut également souligner le support de mes camarades du Latis, notamment George, Thanh Dung, Nicolas et Oussema, avec qui j'ai vécu de bons moments accompagnés par des discussions très intéressantes ; être entouré par des personnes assez compétentes et talentueuses m'a permis d'améliorer sans cesse mon travail.

Estimation de la pression partielle artérielle du dioxyde de carbone par des paramètres noninvasifs

Jihad EL TANNOURY

RÉSUMÉ

Les patients qui souffrent d'un certain trouble respiratoire sont généralement mécaniquement ventilés. Cette ventilation mécanique sert à assurer le dioxygène nécessaire au bon fonctionnement des organes, ainsi qu'à éliminer le dioxyde de carbone produit par les organes. Pour pouvoir bien choisir les paramètres du ventilateur, les cliniciens surveillent les signes vitaux et les tensions artérielles des patients. Obtenir les tensions artérielles directement des artères est douloureux et nuisible aux vaisseaux et les nerfs se trouvant dans l'entourage. Par conséquent, on cherche des méthodes non dangereuses et fiables pour estimer les tensions artérielles. Les tensions de l'oxygène (PaO_2) sont facilement estimées par la saturation transcutanée de ce dernier (SpO_2) obtenue par un spiromètre, alors que l'estimation des tensions du dioxyde de carbone (PaCO_2) est beaucoup plus complexe.

L'objectif principal de travail est de trouver une technique non invasive pour estimer la PaCO_2 d'une manière rapide et fiable dont la performance dépasse celle des méthodes déjà utilisées en pratique comme l'estimation utilisant le volume minute (V_{min}) et celle utilisant le gaz carbonique à la fin de l'expiration (PetCO_2). Pour ce faire, la base de données implémentée dans l'hôpital Sainte Justine dès mai 2015 a été utilisée. Cette base renferme toutes les données patients dans les soins intensifs pédiatriques (PICU). Le Multi Layer Perceptron (MLP), un type de réseau de neurones, a été utilisé pour faire les estimations. L'estimateur utilise un test de gaz sanguins réalisé auparavant pour être calibré. Les données ont été réparties sur quatre groupes : [0h, 2h] ; [2h, 6h] ; [6h, 12h] et [12h, 24h] selon le temps écoulé entre l'acquisition du gaz de calibrage et le gaz à estimer. Pour étudier l'efficacité de l'approche développée, le pourcentage des estimations avec une erreur inférieure à 5 mmHg est adopté puisqu'une erreur inférieure à 5 mmHg a été considérée médicalement comme non significative. Les résultats obtenus montrent que l'approche suivie dans ce travail dépasse les techniques traditionnelles utilisées maintenant en clinique et peuvent améliorer la pratique clinique.

Mots-clés: Détermination de la pression partielle du dioxyde de carbone, Surveillance des Gaz sanguins, MLP, Techniques d'Estimation, Diagnostique

Arterial partial pressures of carbon dioxide estimation using non-invasive parameters

Jihad EL TANNOURY

ABSTRACT

Patients suffering from a respiratory disorder are in general mechanically ventilated. The purpose of this ventilation is to provide the oxygen to the organs to insure their proper function, and to eliminate the carbon dioxide produced by these organs. In order to efficiently choose the ventilator settings, the clinicians rely on vital signs and arterial partial pressure of the patients. Acquiring the arterial partial pressure directly from the arteries is very painful and can destroy the blood vessels and the nerves in the area of the puncture. Therefore, we try to find a reliable and less dangerous way to estimate these partial pressures. The arterial oxygen partial pressure (PaO_2) is easily estimated by the transcutaneous oxygen saturation (SpO_2) measured with a spirometer, while the arterial blood CO_2 pressure (PaCO_2) is more challenging to estimate.

The main purpose of this work is to find a noninvasive reliable technique to rapidly estimate the PaCO_2 with a better performance than the already in-use techniques such as the PaCO_2 estimation with end-tidal CO_2 (PetCO_2) and the minute ventilation volume (V_{min}) evolution. To do so, a high-resolution database in Sainte-Justine University Hospital was used. This database has all patients' data in the Pediatric Intensive Care Unit (PICU). The MultiLayer Perceptron (MLP), a type of the neural networks, was used to do the estimations. The model is calibrated by a previously performed arterial blood gas test. The data were separated into four groups : [0h, 2h], [2h, 6h], [6h, 12h] and [12h, 24h] depending on the time gap between the calibrating test and the time of estimation. To test the efficiency of the developed approach, the percentage of estimations with an error less than 5 mmHg was used since a difference of less than 5 mmHg is considered medically insignificant. The obtained results show that the approach developed in this work is better than the traditional techniques that are used in clinical practice nowadays.

Keywords: Carbon dioxide partial pressure determination, Blood Gas Monitoring, MLP, Estimation techniques, Diagnoses

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 REVUE DE LITTÉRATURE	7
1.1 Axe 1 : Méthodes médicales pour approximer la PaCO ₂	7
1.1.1 La capnographie - PetCO ₂	7
1.1.2 PCO ₂ transcutané (PtcCO ₂)	10
1.1.3 Estimation continue du PaCO ₂ dans la circulation extracorporelle	10
1.1.4 Estimation du PaCO ₂ par l'évolution du volume minute Vmin	12
1.2 Axe 2 : Méthodes mathématiques d'approximation	12
CHAPITRE 2 RÉSEAU DE NEURONES ARTIFICIELS	15
2.1 MLP	15
2.2 Techniques d'évaluation de la performance du réseau	20
CHAPITRE 3 ARTERIAL PARTIAL PRESSURES OF CARBON DIOXIDE ESTIMATION USING NON-INVASIVE PARAMETERS	23
3.1 Introduction	24
3.2 Materials and methods	26
3.2.1 Data	26
3.2.2 Datasets preparation	27
3.2.3 Multi-Layer Perceptron	28
3.2.4 Traditional Techniques	29
3.2.5 Statistical Analyses	30
3.3 Results	31
3.3.1 Variables selection :	33
3.4 Discussion	41
3.5 Conclusion	45
3.6 Acknowledgement	45
CONCLUSION	47
RECOMMANDATIONS	49
ANNEXE I RÉSULTATS DES 5 FOLDS POUR LES MODÈLES MLP	51
ANNEXE II RÉSULTATS DES 5 FOLDS POUR LES MODÈLES MLP + PETCO ₂	57
ANNEXE III LES MEILLEURES COMBINAISONS SELON LES DIFFÉRENTS SOUS GROUPES DANS LA CATÉGORIE 2H	63

BIBLIOGRAPHIE 65

LISTE DES TABLEAUX

	Page
Table 3.1	Patients characteristics - Median and interquartile range 31
Table 3.2	Prediction validation 35
Table 3.3	Accuracy (≤ 5 mmHg) robustness comparison when significant changes occur between calibration ABG and predicted PaCO ₂ 36
Table 3.4	Accuracy (< 5 mmHg) comparison in hypo-, normo- and hypercapnia..... 37
Table 3.5	Bland Altman Characteristics 41

LISTE DES FIGURES

	Page
Figure 0.1	Le parcours des gaz dans le corps 2
Figure 0.2	Les pressions partielles du CO ₂ et du O ₂ dans les veines et artères 3
Figure 1.1	Capnogramme d'un cycle respiratoire 8
Figure 1.2	CDI durant la circulation extracorporelle 11
Figure 2.1	Un réseau MLP de trois couches 16
Figure 3.1	Overall accuracy of the predictive models 34
Figure 3.2	Bland Altman plot of the MLP technique in the 6h category 38
Figure 3.3	Bland Altman plots of the MLP +PetCO ₂ technique in the 6h category (top) and the PetCO ₂ approximation - equation 3.1 (bottom) 39
Figure 3.4	Bland Altman plot of the Vmin approximation in the 6h category - equation 3.2 40

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

ABG	Arterial Blood Gas
CDSS	Clinical Decision Support Systems
CE	Cross Entropy
CEC	Circulation ExtraCorporelle
CO ₂	Dioxyde de carbone
CPAP	Continuous Positive Airway Pressure
ÉTS	École de Technologie Supérieure
ICU	Intensive Care Unit
KPI _v	Indice de capnographie
LATIS	LAaboratoire de Traitement de l'Information en Santé.
MLP	Multi-Layer Perceptron
O ₂	Dioxygène
PICU	Pediatric Intensive Care Unit
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
SMIV	Synchronized Intermittend Mandatory Ventilation

LISTE DES SYMBOLES ET UNITÉS DE MESURE

FiO_2	inspired oxygen Fraction exprimée en pourcentage (%)
HCO_3^-	bicarbonates artériel exprimé en millimole par litre (mmol/L)
ICC	Intra-Class Correlation
log	logarithme népérien
MAwP	Mean Airway Pressure exprimée en centimètre d'eau (cmH ₂ O)
MedAE	Median Average Error
MSE	Mean Squared Error
$PaCO_2$	Pression Artérielle du CO ₂ exprimée en millimètre de mercure (mmHg)
PaO_2	Pression Artérielle du O ₂ (mmHg)
PEEP	Positive End Expiratory Pressure (cmH ₂ O)
$PetCO_2$	End Tidal CO ₂ Pressure (mmHg)
$PexCO_2$	oxygenator EXhaust CO ₂ Pressure (mmHg)
PIP	Positive inspiratory pressure (cmH ₂ O)
$PtcCO_2$	Pression TransCutanée du CO ₂ (mmHg)
RR	respiratory rate exprimé par minute (/min)
SO_2	arterial Oxygen Saturation (%)
SpO_2	transcutaneous Oxygen Saturation (%)
T	température exprimée en °C
V_D	Dead Space exprimé en millilitre (mL)
Vmin	minute Volume exprimé en litre par minute (L/min)
V_T, V_{te}	Tidal Volume exprimé en millilitre par kilogramme (mL/Kg)
$x_{j,i}$	la ième entrée de la jème variable
y_i	la ième sortie

β_i un scalaire

Quelques Conversions

1 cmH₂O = 98 Pascal (Pa)

1 mmHg = 133,322 Pa

1 Pa = 1 Newton/m² (N/m²)

Dans le réseau de neurones

a_k^o	la préactivation de la sortie y_k
b^h, b^o	les biais de la couche cachée et la couche des sorties
f^h, f^o	les fonctions d'activation de la couche cachée et la couche des sorties.
h_j	le j ème neurone de la couche cachée
L_1	la norme 1
L_2	la norme euclidienne
n_j, n_{j+1}	les nombres de variables entrant et sortant d'une certaine couche.
r	coefficient de corrélation Pearson
$U[-a; a]$	une distribution uniforme dans l'intervalle $(-a, a)$
W	l'ensemble des poids
w_{ij}^h	le poids reliant l'entrée x_i avec le neurone h_j de la couche cachée
w_{jk}^o	le poids reliant le neurone h_j de la couche cachée avec la sortie y_k
x_i	la i ème entrée
y_k	la k ème sortie
α, λ	des scalaires
Θ	l'ensemble des hyperparamètres (les poids et les biais)

Θ_i, Θ_{i+1}	les hyperparametres avant et après de leur mise à jour.
∇	le gradient
Ω	un terme de régularisation

INTRODUCTION

Notre travail fait partie d'un très grand projet intitulé Systèmes d'Aide à la Décision Clinique ("Clinical Decision Support Systems") (CDSS) développé dans les soins intensifs (ICU) pédiatriques à l'hôpital Sainte Justine à Montréal. Ce projet vise à améliorer l'assistance clinique aux patients en concevant des outils mathématiques et informatiques intégrés dans les machines déjà utilisées dans les soins intensifs et dans de nouvelles machines à apporter. Plusieurs travaux sont développés dans la création d'une base de données, le transport des données médicales hors de l'hôpital, dans la ventilation mécanique des patients, leur conscience (en analysant les yeux des patients), la détection des arrêts cardiaques à partir des narratives cliniques, reconstruction pulmonaire et beaucoup d'autres. Le travail que nous avons réalisé tombe dans le domaine de la ventilation mécanique, il sert à réduire la durée d'assistance respiratoire et à rendre le traitement plus confortable aux enfants mécaniquement ventilé. Ceci est fait en estimant la pression partielle artérielle du dioxyde de carbone (PaCO_2) à partir des paramètres non invasifs mesurés dans les ICUs. Pour mieux comprendre de quoi il s'agit, on va commencer par expliquer un peu le fonctionnement du système respiratoire avant d'entrer en détails de notre travail.

Le système respiratoire est le système responsable des échanges gazeux au niveau du corps, donc il assure le dioxygène (O_2) nécessaire au bon fonctionnement des organes et expulse le dioxyde de carbone (CO_2) produit vers l'extérieure du corps. Le chemin des gaz dans le corps est présenté par la figure 0.1. La présence de ces gaz dans le sang cause une certaine pression partielle (aussi appelée tension) sur les parois des vaisseaux par chacun des gaz. Plus la quantité du gaz est grande, plus la pression partielle sera grande. Généralement, et si les échanges gazeux se passent tranquillement et sans aucune perturbation, la pression dans les artères exercée par le CO_2 doit être entre 35 et 45 millimètres de mercure (mmHg) alors que celle exercée par l' O_2 doit être entre 75 et 100 mmHg. La figure 0.2 montre les pressions de ces deux gaz dans les artères (en rouge) et les veines (en bleu) au cours des échanges gazeux.

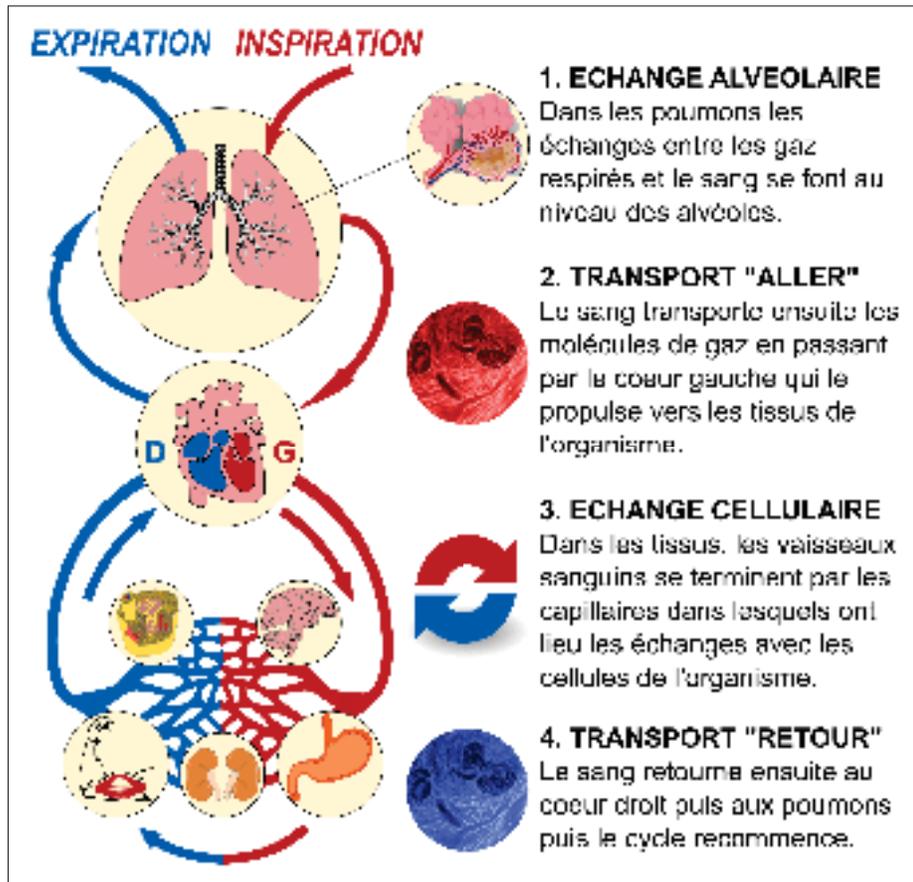


Figure 0.1 Le parcours des gaz dans le corps
Tirée de Passion Subaquatique (2017)

Comme tous les autres systèmes de l'organisme, un mal fonctionnement du système respiratoire peut aboutir à un trouble dans le corps et parfois, dans des cas critiques, sa mort. Lorsque le corps n'est pas capable de rétablir la respiration par lui-même, le patient est soumis à un système d'aide à la respiration qui est souvent un ventilateur, on dit alors que le patient est sous une ventilation mécanique. Le ventilateur sert à assurer la respiration nécessaire au corps en comblant l'insuffisance respiratoire ou bien en assurant une respiration totale au patient. Cette situation persiste jusqu'à ce que les cliniciens trouvent et réparent la cause du mal fonctionnement ou bien le corps se stabilise par lui-même. Bref, surveiller les pressions partielles des gaz peut donner une indication significative sur les échanges et l'état gazeux, ainsi que l'acidité du sang

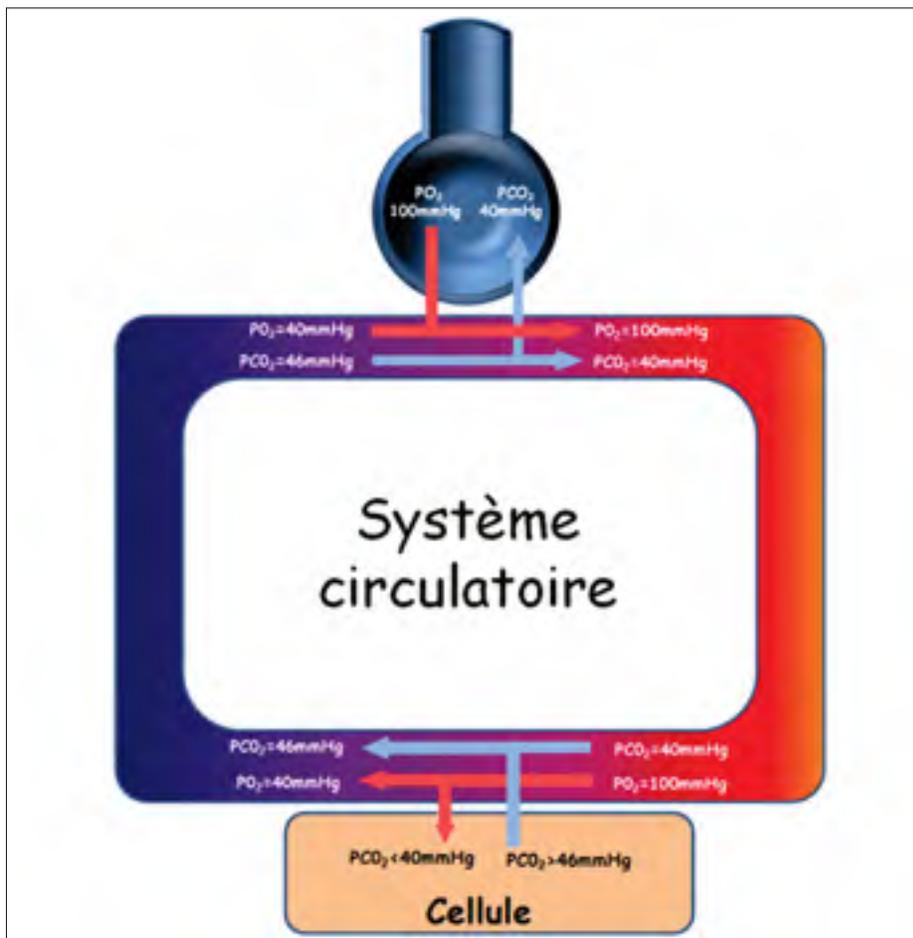


Figure 0.2 Les pressions partielles du CO_2 et du O_2 dans les veines et artères
Tirée du Unisciel (2016)

et de la situation des poumons qui sont les responsables principaux des échanges gazeux. Cela est montré aussi par Sood, Paul & Puri (2010) dans leur étude "Interpretation of arterial blood gas". De plus, ces tensions artérielles sont utilisées pour ajuster les paramètres du ventilateur afin d'assurer une ventilation mécanique efficace et de régler la situation du patient le plus vite possible.

Pour connaître les tensions artérielles, on fait un test des gaz sanguins (ABG : "Arterial Blood Gas"). Ce test consiste à enlever un peu de sang d'une artère à l'aide d'une ponction ou bien

d'un cathéter préexistant et puis envoyer le prélèvement vers un analyseur pour récupérer les mesures. Le test ABG fournit des informations sur les paramètres suivants :

- **pH** : une mesure pour identifier l'état acido-basique du sang. Un pH égal à 7 indique un milieu neutre. Plus ça diminue, plus ça devient acide et plus ça augmente plus ça devient basique. Le pH doit être entre 7.35 et 7.45 chez les personnes saines.
- **PaCO₂** : Pression partielle du dioxyde de carbone dans le sang artériel. Dans le cas normal, cette mesure devrait être entre 35 et 45 mmHg.
- **PaO₂** : Pression partielle du dioxygène dans le sang artériel. Dans le cas normal, il faut avoir une PaO₂ entre 75 et 100 mmHg.

Ces 2 mesures, PaCO₂ et PaO₂, indiquent comment le CO₂ et l'O₂ se déplacent entre le sang et les poumons. Une pression supérieure à la marge normale indique un excès de gaz dans le sang et une pression inférieure à la marge normale indique un défaut. Dans les deux cas, il faut rétablir la situation.

- **SO₂** : la Saturation du sang artériel en oxygène indique le taux d'oxygène contenu dans les globules rouges du sang. La marge normale est 95 à 100%.
- **Lactate** : indique la quantité d'acide lactique dans le sang. La marge normale est de 1 à 2.85 millimoles par litre (mmol/L). Il sert à évaluer la situation acido-basique du sang.
- **HCO₃⁻** : Les ions bicarbonates (HCO₃⁻) jouent un rôle très important dans la régulation du pH dans le sang. Contrairement aux autres mesures, ce paramètre est calculé à partir du pH et du PaCO₂. La marge normale est de 22 à 28 mmol/L. Ces trois paramètres sont reliés par la relation suivante :

$$pH = 6,1 + \log \frac{[HCO_3^-]}{[CO_2d]}$$

où $[CO_2d] = a \times PCO_2$ (a = 0,03 mmol/L/mmHg)

Malheureusement, le prélèvement du sang ne peut pas se faire d'une manière continue et peut être très douloureux spécialement chez les enfants ; la procédure peut causer des infections et des

caillots sanguins, et peut aussi endommager les vaisseaux et les nerfs se trouvant à proximité, affirme la World Health Organization et al. (2010). Du coup, les chercheurs ont essayé de trouver des alternatives pour pouvoir mesurer ou approximer ces pressions artérielles d'une manière non invasive.

Les patients mécaniquement ventilés dans les soins intensifs sont aussi surveillés par des moniteurs pour vraiment savoir leur état clinique. Parmi les paramètres non invasifs mesurés au cours de la surveillance, on distingue :

- **RR** : la fréquence respiratoire ("Respiratory Rate") représente le nombre de respirations effectuées par minute. La valeur normale normale de RR est de 12 à 20 respirations par minutes pour les adultes (12 ans et plus) alors que pour les enfants, la marge normale est de 20 à 40 respirations par minute.
- **FiO₂** : la Fraction inspirée en Oxygène qui correspond au pourcentage d'oxygène dans le mélange gazeux inspiré. La fraction normale dans l'air inspiré est de 21%, mais ce pourcentage varie beaucoup durant la ventilation mécanique pour assurer une plus grande quantité d'oxygène aux patients qui souffrent d'une insuffisance respiratoire, elle peut atteindre les 100% aussi.
- **PEEP** : de l'anglais "Positive End Expiratory Pressure", c'est la pression (au-delà de la pression atmosphérique) dans les poumons à la fin de l'expiration. Normalement, la PEEP devrait être entre 3 et 5 cmH₂O
- **PIP** : de l'anglais "Peak Inspiratory Pressure", qui correspond à la pression maximale appliquée sur les poumons durant l'inspiration. Il vaut mieux que la PIP soit entre 16 et 18 cmH₂O et ne dépasse pas 25 cmH₂O
- **MAwP** : "Mean Airway Pressure", c'est la pression de maintien d'ouverture du poumon sur tout le cycle respiratoire.

- **SpO₂** : "Peripheral Oxygen Saturation" aussi connue par "transcutaneous oxygen saturation". La différence entre cette mesure et la SO₂ des ABG c'est que la SpO₂ est mesurée par des capteurs de surface et non pas directement à travers du sang. Les études montrent que dans la majorité des situations, la SpO₂ reflète bien la SO₂. La marge normale du SpO₂ est 96% à 99%.
- **Vte** : "Expiratory Tidal Volume", cette mesure correspond au volume d'air échangé pendant une inspiration ou une expiration normale (non forcée). Le Vte doit être entre 5 et 12 ml/Kg.
- **Vmin** : le Volume Minute représente le volume d'air inspiré ou expiré par les poumons pendant une minute. Ou bien c'est la RR multipliée par le Vte. La valeur normale de ce paramètre varie autour de 0.2L/Kg/min.

Tous ces paramètres sont avérés être en relation avec les pressions artérielles ; la PaO₂ peut être estimée par la mesure du SpO₂, mais l'estimation du PaCO₂ est un peu plus complexe et plusieurs études ont été réalisées pour trouver un moyen d'estimer la PaCO₂. Les deux groupes de paramètres sont utilisés pour faire les ajustements du ventilateur. En effet, les cliniciens se base sur la PaO₂ avec la SpO₂ pour ajuster la PEEP et la FiO₂, alors que le pH et la PaCO₂ sont utilisés pour modifier la PIP, la RR et le Vte.

Dans notre projet, on va utiliser les paramètres non invasifs détailler plus haut pour estimer la PaCO₂ en calibrant l'estimateur par un ABG fait en avance. Dans un premier lieu, on va expliquer les différentes techniques médicales et mathématiques utilisées pour estimer la PaCO₂, ensuite on va détailler le réseau de neurones qu'on a utilisé pour faire notre estimation et puis on présentera l'article qu'on a rédigé dans ce contexte.

CHAPITRE 1

REVUE DE LITTÉRATURE

Notre revue de littérature est basée sur deux axes : Le premier axe concerne les différents alternatifs et méthodes médicales pour estimer la PaCO_2 , c'est-à-dire les différents paramètres et les différentes combinaisons entre eux donnant une approximation du PaCO_2 . Le second axe concerne les modèles mathématiques d'estimation comme la régression linéaire et les méthodes d'apprentissage machine.

1.1 Axe 1 : Méthodes médicales pour approximer la PaCO_2

1.1.1 La capnographie - PetCO_2

La capnographie est une technique de mesure de CO_2 expiré en fonction du temps généralement utilisée dans les anesthésies et les soins intensifs. Un émetteur envoie des ondes infrarouges vers le receveur à travers le flux gazeux pour savoir l'état du gaz expiré. Les chercheurs ont eu l'idée d'utiliser cette information pour étudier les échanges gazeux à l'intérieur du corps. Alors ils ont commencé par approximer la PaCO_2 à partir du CO_2 expiré. Par la capnographie on mesure le CO_2 de manière continue, mais ce qui nous importe c'est la mesure à la fin de l'expiration. Cette mesure est appelée "end tidal CO_2 pressure" (PetCO_2) et généralement elle sous-estime la PaCO_2 . La figure 1.1 représente un capnogramme d'un cycle respiratoire.

Plusieurs recherches ont été réalisées pour étudier l'efficacité d'utiliser la PetCO_2 pour estimer la PaCO_2 , la corrélation entre la PetCO_2 et la PaCO_2 paraît être intacte face à un changement de température corporelle. Brothers, Ganio, Hubing, Hastings & Crandall (2011) ont étudié la corrélation chez un groupe de patients avec un stress thermique modéré et un autre avec un stress thermique sévère. Le changement de température a affecté la PaCO_2 et la PetCO_2 de la même manière et la corrélation n'a pas changé. Mais il semble que la corrélation est largement dépendante de beaucoup d'autres conditions médicales du patient. McSwain, Hamel, Smith, Gentile, Srinivasan, Meliones & Cheifetz (2010) ont montré que la corrélation

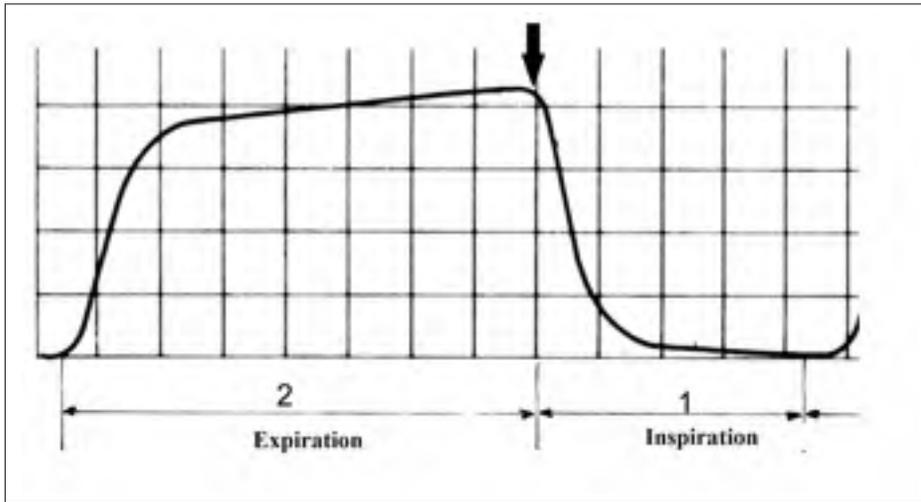


Figure 1.1 Capnogramme d'un cycle respiratoire

entre la $PaCO_2$ et la $PetCO_2$ sont inversement proportionnelles au rapport espace mort sur le volume tidal V_D/V_T . L'espace mort connu par le "Dead Space" constitue le volume d'air qui ne participe pas dans les échanges gazeux alors que le volume tidal constitue le volume d'air inspiré ou expiré normalement sans effort supplémentaire. La corrélation a diminué de 0,95 pour $V_D/V_T \leq 0,40$ à 0,78 pour $V_D/V_T > 0,7$. Dans une autre étude reliant la sévérité des maladies pulmonaires à la corrélation $PaCO_2$ - $PetCO_2$, Panigrahi, Ahmed, Mishra & V (2018) ont prouvé que la corrélation diminue lorsque la situation devient plus sévère. La sévérité des maladies était évaluée par le rapport PaO_2/FiO_2 , pour les patients oxygénés normalement ayant un rapport $PaO_2/FiO_2 > 200$ la corrélation était plus élevée que celle chez les patients hypoxiques ayant un rapport $PaO_2/FiO_2 < 200$ (0,93 et 0,86 respectivement). De plus, le type de ventilation a aussi un effet sur la relation. Razi, Moosavi, Omidi, Saebi & Razi (2012) ont étudié ceci suivant trois modes de ventilation : ventilation invasive (SMIV : "Synchronized Intermittend Mandatory Ventilation"), ventilation non invasive (CPAP : "Continuous Positive Airway Pressure") et le T-Tubing. La corrélation la plus élevée était dans le mode T-Tubing ($r=0,92$) puis dans le mode invasif ($r=0,89$) et la moins élevée dans le mode non invasive ($r=0,84$).

L'efficacité du $PetCO_2$ dans l'estimation du $PaCO_2$ dépend aussi de la nature de la capnographie ; Chan, Chan & Gin (2003) ont comparé les résultats obtenus par deux types de capnographies : la capnographie mainstream et la capnographie sidestream. La différence entre ces deux types c'est que dans la première, la $PetCO_2$ est mesurée sur le tuyau principal alors que dans le second, la $PetCO_2$ est mesurée sur un échantillon pris du tuyau principal. Le mainstream a apporté de meilleurs résultats que le sidestream avec une différence de 0,64 kiloPascal (kPa) et 0,99 kPa respectivement.

Pour aller plus loin, Baudin, Bourgoïn, Brossier, Essouri, Emeriaud, Wysocki & Jouvét (2016) ont essayé d'améliorer l'estimation en utilisant une forme corrigée de $PetCO_2$. Cette formulation repose sur deux autres mesures noninvasives : 1) la $MAwP$ et 2) la FiO_2 . La régression multiple leur a donné la relation suivante :

$$PaCO_2 = -1.525 + 0.876 \times PetCO_2 + 0.442 \times MAwP + 0.117 \times FiO_2 \quad (1.1)$$

De plus, les chercheurs ont amélioré davantage cette forme corrigée en ajoutant l'index de capnographie $KPIv$ (ce sont des mesures tirée du capnogramme) :

$$PaCO_2 = 0.859 + 0.827 \times PetCO_2 + 0.310 \times MAwP + 0.081 \times FiO_2 + 0.529 \times KPIv \quad (1.2)$$

En utilisant la deuxième forme corrigée, Baudin *et al.* (2016) ont pu augmenter le pourcentage des erreurs moins que 5 mmHg de 75% à 95% avec un $R^2 = 0,90$. Dans notre étude, n'ayant pas accès aux capnogrammes, l'équation 1.1 a été utilisée pour comparer avec notre méthode.

1.1.2 PCO₂ transcutané (PtcCO₂)

Une autre façon d'estimer la PaCO₂ c'est la PCO₂ transcutané. Cette méthode consiste à mesurer la diffusion du CO₂ à travers les tissus du corps et la peau à l'aide d'un capteur de surface. Suite au développement technologique, un capteur de surface peut mesurer la PtcCO₂ ou bien à la fois la PtcCO₂ avec la PaO₂ ou la SpO₂. Les techniques transcutanées sont plus fiables chez les enfants et les nouveaux nés que chez les adultes puisque leur peau est mince et très vascularisée. Selon Eberhard (2007), la PtcCO₂ est supérieure à la PaCO₂ en principe puisque : 1) la peau produit elle-même du CO₂ qui va s'ajouter au CO₂ diffusé, et 2) l'échauffement du capteur cause un apport supplémentaire de CO₂. Le métabolisme de la peau ajoute environ 5 mmHg de pression alors que l'élévation de température du capteur ajoute 4.5% par °C. Pour ces raisons, les fabricants de systèmes de PtcCO₂ incluent des corrections à leur système. Lermuzeaux, Meric, Sauneuf, Girard, Normand, Lofaso & Terzi (2016) ont fait une étude comparative entre la PetCO₂ et la PtcCO₂ et leur efficacité dans l'estimation du PaCO₂. La corrélation PtcCO₂-PaCO₂ était supérieure à la corrélation PetCO₂-PaCO₂ chez les patients adultes non intubés dans les soins intensifs qui souffrent de troubles respiratoires sévères (r=0,97 et r=0,62 respectivement).

1.1.3 Estimation continue du PaCO₂ dans la circulation extracorporelle

Parmi les techniques qui estiment le PaCO₂ d'une manière continue, on distingue le "CDI Blood Parameter monitoring system 500 " ; c'est un équipement utilisé uniquement durant la circulation extracorporelle (CEC) et n'est pas applicable ailleurs. La figure 1.2 représente le principe de cette technique. Un capteur est mis sur la voie veineuse et un autre sur la voie artérielle pour mesurer les pressions veineuses et artérielles. Des émetteurs envoient des ondes excitatrices vers le sang avec des longueurs d'onde différentes, ces ondes vont interagir avec les molécules du sang et une lumière sera émise. Ensuite, le système capte cette lumière et donne une mesure (pression et pH) dépendamment de l'intensité de la lumière émise ; plus l'intensité est élevée signifie un plus grand nombre de particules et alors une mesure élevée. Cette technique est assez rapide et peut donner une estimation du pH, du PCO₂ et du PO₂ à chaque seconde. Gelsomino et al. (2011) ont fait une étude sur 10 porcs pour tester l'efficacité de cette technique, ils ont

trouvé une corrélation $r > 0.97$ pour le pH, le PCO_2 , le PO_2 et le SO_2 .

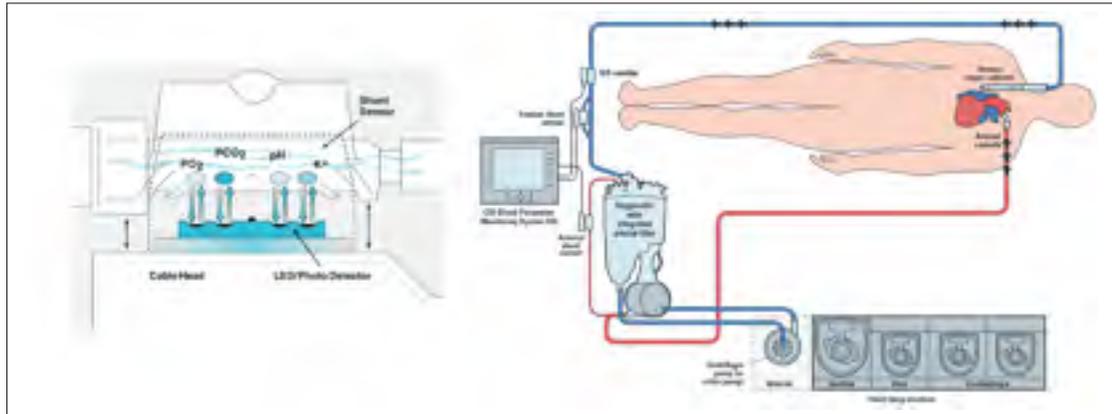


Figure 1.2 CDI durant la circulation extracorporelle
Tirée de Terumo CDI500 (2018)

Une autre technique d'estimation continue c'est la capnographie par oxygénateur d'échappement ("oxygenator exhaust capnography"). Durant la circulation extracorporelle, le sang circule à travers un oxygénateur dans lequel la PCO_2 sera mesurée, cette pression sera notée $PexCO_2$. Potger, McMillan, Southwell, Dando & O Shaughnessy (2003) ont trouvé une bonne corrélation $r = 0.91$ entre la $PaCO_2$ et le $PexCO_2$. Dans le même contexte, la corrélation entre la $PaCO_2$ et le $PexCO_2$ pour des patients normothermiques (ayant une température normale de $37^{\circ}C$) était $r = 0.79$ et pour des patients hypothermiques $r = 0.78$ durant une étude réalisée par Baraka, El-Khatib, Muallem, Jamal, Haroun-Bizri & Aouad (2005).

Ces techniques paraissent être efficaces dans l'estimation continue du $PaCO_2$, mais malheureusement, elles ne peuvent pas être applicables dans tous les cas cliniques puisqu'elles nécessitent une voie sanguine extracorporelle (rarement trouvée en clinique) pour les appliquer.

1.1.4 Estimation du PaCO₂ par l'évolution du volume minute Vmin

Cette approche est la plus ancienne et la plus simple pour estimer la PaCO₂. C'est une relation entre la PaCO₂ et le volume minute mesurés à un instant t et un instant $t + 1$.

$$PaCO_{2t+1} = \frac{PaCO_{2t} \times Vmin_t}{Vmin_{t+1}} \quad (1.3)$$

Pour que cette technique soit efficace, Wexler & Lok (1981) ont montré que ces trois conditions doivent être satisfaites, sinon, l'estimation ne devient plus précise :

1. le patient doit être dans un état métabolique stable durant la production du dioxyde de carbone.
2. le rapport espace mort sur le volume tidal V_D/V_T doit rester inchangé, le changement au niveau du Vmin doit se faire en variant la fréquence respiratoire et non pas le volume tidal.
3. le patient doit accepter le nouveau volume minute sans outrepasser le ventilateur.

Ces conditions sont vraiment strictes et rarement on peut trouver de telles situations. Pourtant, les cliniciens appliquent la relation 1.3 pour estimer la PaCO₂. En plus de la forme corrigée de la PetCO₂, on a comparé notre méthode à l'approche utilisant le Vmin.

1.2 Axe 2 : Méthodes mathématiques d'approximation

La majorité des recherches dans le domaine de notre étude c'est-à-dire les recherches qui visent trouver une relation entre un ou plusieurs paramètres et la PaCO₂ sont des études réalisées par la régression linéaire simple ou multiple. Rares sont les recherches qui utilisent d'autres techniques comme les réseaux neuronaux et l'apprentissage machine. Dans un premier lieu, on va expliquer un peu la régression linéaire puis on va comparer cette technique avec les réseaux MultiLayer Perceptron (MLP).

La régression linéaire est une méthode par laquelle on essaie de trouver une relation linéaire entre les variables, si on a juste deux variables alors on parle de la régression linéaire simple alors que si on a plus que deux variables, la régression devient multiple. En d'autres termes, si on veut estimer une variable Y (ensemble de y_i) à partir de plusieurs paramètres X_j dont chacun est formé d'un ensemble de $x_{j,i}$ par la régression linéaire, le problème s'écrit comme suit :

$$y_i = \beta_0 + \beta_1 \times x_{1,i} + \beta_2 \times x_{2,i} + \dots + \beta_n \times x_{n,i} \quad (1.4)$$

Où β_i sont des scalaires.

Le problème de cette technique est sa linéarité. La régression linéaire a des limitations là où la relation devient non linéaire entre les variables. Plusieurs études ont été réalisées pour comparer la régression linéaire et des techniques d'apprentissage machine. Martinez-Coll & Nguyen (2001) ont comparé la régression multiple avec les réseaux de neurones dans l'estimation des mesures corticales ; la variance des valeurs obtenues par la régression était 17,57 alors que celle des valeurs obtenues par les réseaux était 14,1. Shan & Jie (2011) ont aussi fait une étude comparative entre les deux techniques pour estimer la cyclosporine, l'apprentissage machine a donné un pourcentage de 90% de bonnes estimations alors que la régression a donné un pourcentage de 88%. Une autre étude réalisée par Bhaduri, Bhaduri, Bhaduri & Mohapatra (2009) montre aussi que les réseaux de neurones artificiels sont plus efficaces que la régression linéaire dans la prédiction de la pression sanguine ; la corrélation obtenue par la régression était 0,65 alors que celle obtenue par les réseaux de neurones était 0,82. Toujours dans le domaine médical, une étude pour prédire les hémorragies intracrâniennes réalisée par Zernikow, Holtmannspoetter, Michel, Theilhaber, Pielemeier & Hennecke (1998) montre aussi que l'apprentissage machine est mieux que la régression, la surface sous la courbe ROC du réseau était 0,94 alors que celle de la régression est 0,88. (La courbe ROC est une courbe qui représente le taux des vrais positifs en fonction du taux des faux positifs pour évaluer la performance des classificateurs.)

Dans d'autres domaines, l'apprentissage machine a aussi montré qu'il peut être efficace dans la prédiction. Dans l'environnement par exemple, Adel Zaqoot, Baloch, Khaliq Ansari & Ali Unar (2010) ont trouvé que dans la prédiction du pH, le modèle basé sur les MLP a donné une corrélation de 0,86 largement supérieure à celle obtenue par la régression (0,15). Dans le même domaine aussi et avec le même résultat obtenu, une étude réalisée par Choubin, Khalighi-Sigaroodi, Malekian & Kişi (2016) montre ceci. Dans le domaine de finance, les algorithmes d'apprentissage machine peuvent donner des très bons résultats, dans leur étude, Naeini, Taremian & Hashemi (2010) montrent l'efficacité des MLPs dans la prédiction des stocks : les erreurs obtenues par les MLPs étaient au alentour de 10% alors que celles obtenues par la régression linéaire est de 20%.

CHAPITRE 2

RÉSEAU DE NEURONES ARTIFICIELS

L'idée des réseaux de neurones artificiels a été inspirée au départ par le système nerveux (McCulloch & Pitts (1943)). Les chercheurs ont réussi à modéliser le système nerveux par des circuits électriques. À l'époque, les ordinateurs n'étaient pas assez puissants pour pouvoir simuler ces réseaux artificiels, mais avec l'avancement technologique vers les 1950s, on a pu réaliser cette simulation. En développant ces réseaux, on a constaté qu'ils sont capables de trouver des liens linéaires et non linéaires entre les données. Cette propriété très intéressante semblait être utile dans la modélisation d'autres situations et non pas juste le système nerveux comme les estimations, la prédiction et la classification des données dans n'importe quel domaine. Plusieurs types de réseaux de neurones ont été créés et probablement le plus populaire c'est le MultiLayer Perceptron (MLP).

2.1 MLP

Les MLPs de Gardner & Dorling (1998) sont des réseaux neuronaux "feedforward" c'est-à-dire que les connexions entre les neurones ne forment pas de cycle ; chaque couche communique uniquement avec la couche suivante. Ce type de réseaux est très puissant dans la modélisation des problèmes complexes. Les MLPs sont généralement composés d'au moins trois couches : 1) la couche des entrées, 2) la couche cachée et 3) la couche des sorties. Dans plusieurs applications, on peut ajouter deux, trois ou même plus de couches cachées au réseau dans le but d'améliorer les résultats de la prédiction ou de la classification. Chaque neurone de la couche cachée et de la couche de sortie est connecté à tous les neurones de la couche précédente. L'entraînement du réseau est fait d'une manière supervisée c'est-à-dire que le modèle apprend des exemples d'entrées et des sorties correspondantes durant la phase d'apprentissage. Tout d'abord, on fait passer les données dans le réseau, les neurones de la couche des entrées prennent alors les valeurs de ces données. Les couches cachées sont responsables du traitement de l'information et de la sélection des fonctionnalités (features selection) qui semblent être importantes dans

l'estimation, classification ou prédiction des données. Et finalement, on obtient les estimations ou les classifications par la couche des sorties. Pour passer d'une couche à une autre, on calcule une préactivation et puis une activation. La préactivation est une combinaison linéaire des neurones, les poids et le biais. Le rôle du biais est similaire à l'ordonnée à l'origine dans l'équation d'une droite. La fonction d'activation est une fonction non linéaire qu'on l'applique sur le résultat de la préactivation. Pour mieux comprendre ce type de réseau de neurones et comment il fonctionne, on a illustré dans la figure 2.1 un réseau MLP de trois couches avec n entrées, m neurones dans la couche cachée et o sorties.

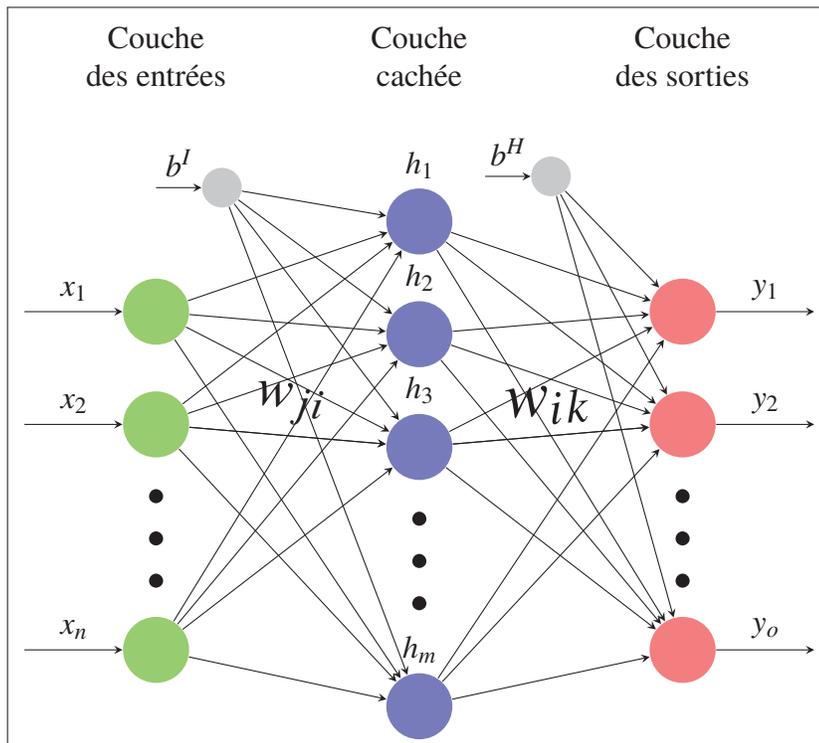


Figure 2.1 Un réseau MLP de trois couches

Si on note les entrées par des x_i , $i \in [1 \dots n]$, les sorties par des y_k , $k \in [1 \dots o]$, les neurones de la couche cachée par h_j , $j \in [1 \dots m]$, le poids reliant un neurone h_j de la couche cachée à un neurone y_k de la couche des sorties par w_{jk}^o et w_{ij}^h le poids reliant le neurone x_i de l'entrée avec le neurone h_j de la couche cachée, la préactivation du neurone y_k de la couche des sorties par a_k^o , les biais des couche cachée et couche des sorties par b^h et b^o respectivement et la fonction

d'activation de la couche cachée et la couche des sorties par f^h et f^o respectivement ; alors la propagation en avant (forward propagation) au niveau du réseau sera calculée de la manière suivante :

$$a_k^o = b^o + \sum_{j=1}^m w_{jk}^o y_j^h \quad (2.1)$$

$$y_k^o = f^o(a_k^o) \quad (2.2)$$

$$y_k^o = f^o\left(b^o + \sum_{j=1}^m w_{jk}^o f^h\left(b^h + \sum_{i=1}^n w_{ij}^h x_i\right)\right) \quad (2.3)$$

$$k = 1 \dots o$$

Les fonctions d'activation f les plus populaires dans le domaine d'apprentissage machine sont :

- la fonction sigmoïde : $\sigma(x) = \frac{1}{1+e^{(-x)}}$,
- la fonction tangente hyperbolique : $\tanh(x) = \frac{e^{2x}-1}{e^{2x}+1}$,
- la fonction "Rectified linear unit" (ReLU) : $\text{ReLU}(x) = \max(0, x)$.

La fonction sigmoïde souffre d'une convergence lente et de la "disparition du gradient" (vanishing gradient) ; ce phénomène prend place lorsque le gradient durant la rétropropagation devient trop petit et tend vers zéro. La fonction tangente hyperbolique souffre aussi de ce problème de "disparition du gradient". Dans notre réseau, on va utiliser la fonction *ReLU*. Cette fonction est devenue très populaire récemment puisqu'elle peut aboutir à une convergence du gradient six fois plus rapide que la fonction *tanh*.

Pour initialiser les poids, certains utilisent une distribution normale peut attribuer des valeurs à partir de cette distribution. D'autres prennent des valeurs aléatoires entre 0 et 1. Il existe d'autres méthodes aussi, mais dans notre travail, on a utilisé l'initialisation de Glorot & Bengio (2010) décrite dans leur article, nommée initialisation de Xavier ou bien initialisation de Glorot définie

par :

$$W \sim U\left[-\frac{\sqrt{6}}{\sqrt{(n_j + n_{j+1})}}; \frac{\sqrt{6}}{\sqrt{(n_j + n_{j+1})}}\right] \quad (2.4)$$

Où W représente l'ensemble des poids, $U[-a; a]$ représente une distribution uniforme dans l'intervalle $(-a, a)$, n_j et n_{j+1} sont les nombres d'entrées et de sorties de la couche considérée. Cette initialisation semble être très intéressante puisqu'elle cause une convergence plus rapide du gradient et par suite une phase d'entraînement plus rapide du réseau.

Durant l'apprentissage du réseau (qui est expliqué par la suite), les poids w_{ij} sont mis à jour de façon à réduire les erreurs de prédiction. Les fonctions d'erreurs les plus utilisées dans l'apprentissage machine sont l'erreur quadratique moyenne (MSE : mean squared error) et l'entropie croisée (CE : cross entropy). La CE est généralement utilisée dans les problèmes de classifications, dans notre projet on a utilisé la MSE définie par :

$$l = MSE = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i))^2 \quad (2.5)$$

Où y_i est la valeur correcte et $f(x_i)$ est la valeur prédite calculée par le réseau pour la i^{me} exemple d'apprentissage.

Pour améliorer encore plus la performance de l'apprentissage, et pour lutter contre le surapprentissage, on ajoute un terme de régularisation Ω à l'erreur :

$$L = MSE + \lambda\Omega \quad (2.6)$$

où λ est un paramètre qui contrôle la puissance de la régularisation. Le terme de régularisation est basé sur la norme L_1 et la norme L_2 :

- $\Omega = L_1 = \sum |\mathbf{w}|$

$$- \Omega = L_2 = \sum \mathbf{w}^2$$

Dans notre cas, on a une seule sortie alors la norme L_1 n'est autre que la différence entre la valeur prédite et la valeur réelle et la norme L_2 est le carré de cette différence. Dans la descente du gradient au cours de l'apprentissage, il vaut mieux utiliser la norme L_2 puisque la situation ressemble à une parabole (degré 2) alors que la norme L_1 ressemble à un triangle et peut donner une singularité. Ceci peut causer un problème dans le calcul du gradient et peut bloquer l'apprentissage. Alors, on a utilisé la norme L_2 dans notre réseau.

La rétropropagation (back propagation) expliqué par Rumelhart, Hinton, Williams et al. (1988) consiste à calculer le gradient des erreurs L déjà obtenues par l'équation (2.6) par rapport aux hyperparamètres Θ formés par l'ensemble des poids et des biais dans le réseau :

$$\Delta = -\nabla_{\Theta} l(\Theta) - \lambda \nabla_{\Theta} \Omega(\Theta) \quad (2.7)$$

Puis les hyperparamètres sont mis à jour par l'équation suivante :

$$\Theta_{i+1} = \Theta_i + \alpha \Delta \quad (2.8)$$

Le processus d'entraînement du réseau se déroule de la manière suivante :

- l'apprentissage qui est divisé lui-même en deux parties : l'entraînement et la validation. On fait apprendre le modèle sur un groupe de données, qu'on appelle le groupe d'apprentissage, pour créer un modèle et puis on le valide sur un autre groupe, le groupe de validation, pour éviter le problème de sur-apprentissage (overfitting). Le "overfitting" se produit lorsque le réseau apprend beaucoup sur le groupe d'apprentissage et sera inadéquat de généraliser sur d'autres données comme si le modèle créé est spécifique pour les données utilisées dans l'entraînement uniquement. Pour éviter l'overfitting, il faut arrêter l'apprentissage lorsque la précision de la classification/prédiction augmente sur le groupe d'apprentissage, mais reste la même ou diminue sur le groupe de validation. Durant l'apprentissage, on fait plusieurs

propagations en avant et des rétropropagation jusqu'à ce que la précision n'augmente plus ou bien on risque d'atteindre le sur-apprentissage.

- le test : durant lequel on teste la précision et l'efficacité du réseau à classifier ou prédire à partir des données qu'il n'a jamais utilisées.

Généralement, le groupe d'apprentissage constitue 70% de la totalité des données, le groupe de validation est constitué de 15% des données et les 15% restantes forment le groupe de test.

Pour récapituler, une itération d'entraînement consiste à faire une propagation en avant pour calculer les sorties y_i par les équations 2.1, 2.2 et 2.3, puis on calcule les erreurs L par l'équation 2.6 et finalement on fait la rétropropagation pour mettre à jour les poids par l'équation 2.8. Ceci est répété plusieurs fois sur chaque exemple de l'ensemble d'apprentissage jusqu'à ce que l'erreur de validation devient minimale ou quand on commence à sur-apprendre le réseau.

Et finalement, pour choisir le nombre de couches cachées et le nombre de neurones dans chaque couche, on a essayé plusieurs architecture avec une, deux et trois couches cachées ayant un à trente neurones par couche. L'architecture avec le meilleur résultat a été retenue.

2.2 Techniques d'évaluation de la performance du réseau

Il existe beaucoup de techniques et de méthodes pour évaluer la performance des réseaux neuronaux et ça dépend du type des réseaux, du domaine d'étude et de la nature des données. Les techniques utilisées pour évaluer la classification sont différentes de celles utilisées pour évaluer la régression et l'estimation des données. Dans notre projet, le score principal utilisé pour évaluer les modèles est le pourcentage des estimations avec une erreur $< 5\text{mmHg}$ et $< 10\text{ mm Hg}$ puisque 5 mmHg de différence est considéré comme insignifiant médicalement. En addition, on a utilisé les métriques suivantes :

1. La médiane des erreurs absolues ("Median Absolute Error") *MedAE* :

$$MedAE = \text{median}(|\hat{y}_1 - y_1|, \dots, |\hat{y}_n - y_n|) \quad (2.9)$$

Avec y_i les valeurs réelles des sorties et \hat{y}_i les valeurs estimées.

2. Le coefficient de corrélation Pearson r :

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.10)$$

Où x_i est la valeur réelle, y_i est la valeur prédite, \bar{x} la moyenne des x_i et \bar{y} la moyenne de y_i .

3. La corrélation intraclasse (ICC) pour étudier la fiabilité de l'approche proposée. Plusieurs types de corrélations intraclasse existent, on se base sur comment l'étude a été conduite pour choisir le type (d'après Shrout & Fleiss (1979) et Koo & Li (2016)). Puisqu'on utilise une seule technique pour faire l'estimation de toutes les valeurs, le MLP, et puisqu'on calcule la fiabilité d'une seule prédiction (le MLP donne une seule estimation du PaCO₂ et non pas plusieurs) alors le type utilisé est la ICC (3,1).

On a utilisé les représentations Bland Altman pour visualiser l'accord entre les vraies valeurs du PaCO₂ et les valeurs prédites. Bland & Altman (1986) ont montré que même avec un coefficient de corrélation très élevé, on peut ne pas avoir un accord entre deux distributions. Si on prend par exemple deux droites ayant une même pente et des ordonnées à l'origine différentes, on aura une corrélation maximale ($r = 1$), mais les valeurs de l'une sont différentes de l'autre.

CHAPITRE 3

ARTERIAL PARTIAL PRESSURES OF CARBON DIOXIDE ESTIMATION USING NON-INVASIVE PARAMETERS

Jihad El Tannoury¹, Michael Sauthier MD¹, Philippe Jovet MD PhD¹, Rita Noumeir PhD¹

¹ Département de Génie Électrique, École de Technologie Supérieure,
1100 Notre-Dame Ouest, Montréal, Québec, Canada H3C 1K3

² Service des soins intensifs pédiatriques, Hôpital Sainte Justine,
3175 Chemin de la Côte-Sainte-Catherine, Montréal, QC H3T 1C5

Article soumis pour publication à la revue « IEEE Transactions on Biomedical Engineering » en
Novembre 2019.

ABSTRACT

Objective : We aim to create a predictive model capable of giving a noninvasive, immediate and reliable estimate of the arterial partial pressure of carbon dioxide (PaCO_2) with a better reliability than its estimation from end-tidal CO_2 (PetCO_2) and minute ventilation volume (V_{min}) evolution.

Methods : We collected data from the Intensive Care Unit (ICU) database of Sainte-Justine University Hospital (Montreal, Canada) and used the multilayer perceptron (MLP) to estimate the PaCO_2 . Input data were (1) Arterial blood gas (ABG) at a previous time to calibrate the model, (2) mechanical ventilator parameters and (3) pulse oximetry. The data were divided into four groups depending on the time gap between previous ABG and its prediction : [0h, 2h], [2h, 6h], [6h, 12h] and [12h, 24h].

Results : We included 17,329 ABGs collected from 527 patients between May 2015 and October 2018. Median age was 6.7 months (interquartile range 1-60) and female proportion was 45%. Patients had a median of 13 ABGs per patient (IQR 5-34). The accuracy of the models in the four groups was 18%, 18%, 19% and 25% higher than the minute volume models and the PetCO_2 models (4% to 11%, respectively).

Conclusion : Our model based on noninvasive parameters was able to better estimate the PaCO₂ when compared to the traditional techniques.

Significance : ABG analysis is very important in ICU ; it is the gold standard in respiratory and acid-base evaluation. ABG is invasive, painful and risky. Our approach, noninvasive and reliable, is an alternative for optimizing mechanical ventilator settings, thus providing better care for patients.

Keywords – Carbon dioxide partial pressure determination, Blood Gas Monitoring, MLP, Estimation techniques, Diagnosis.

3.1 Introduction

ARTERIAL blood gas analysis is one of the most common blood analyses for managing patients in intensive care unit (ICU) since it gives important information about the patient's oxygenation, ventilation status, acid-base and gas exchange status Sood *et al.* (2010). This test requires a small amount of blood to be extracted from an artery either from an indwelling catheter (in case of repeated measurements) or by performing a direct arterial puncture. The procedure can be painful, especially in children, and is at-risk of infections, clots, nerve damage or vessel damage Organization *et al.* (2010). Clinicians rely on the ABG results to make changes in the ventilators settings to either improve oxygenation or CO₂ removal. However, the analysis can only be done intermittently. Some continuous estimations have been developed for both oxygenation and CO₂ removal assessment.

The oxygenation (arterial partial pressure of oxygen - PaO₂) is estimated by the continuous monitoring of transcutaneous oxygen saturation (SpO₂) that is reliable to detect hypoxemia. The arterial partial pressure of carbon dioxide (PaCO₂) is more complex to estimate and is very important to measure since it is the reference to assess pulmonary ventilation and it is a critical measurement for a number of diseases seen in ICU Jakkula, Reinikainen, Hästbacka, Loisa, Tiainen, Pettilä, Toppila, Lähde, Bäcklund, Okkonen *et al.* (2018); Vaahersalo, Bendel, Reinikainen, Kurola, Tiainen, Raj, Pettilä, Varpula, Skrifvars, Group *et al.* (2014). The simplest

approach to estimate PaCO₂ uses the minute volume (V_{min}), defined by the exhaled volume of gas per breath time respiratory rate Wexler & Lok (1981). However, this estimation is reliable in strict stable clinical conditions that are nearly impossible to achieve in clinical practice *ie* constant carbon dioxide production and stable dead space (V_D). Capnography is another technique to estimate the PaCO₂. With capnography, we study the exhaled carbon dioxide by measuring the end tidal carbon dioxide pressure (PetCO₂) and we make inferences about the respiration conditions of the patient and his PaCO₂. These two measures are highly correlated in healthy individuals. But patients with severe lung diseases have a high dead space fraction, this negatively impacts the correlation between PaCO₂ and PetCO₂ McSwain *et al.* (2010); Panigrahi *et al.* (2018); Wexler & Lok (1981). The ventilation type (invasive, non-invasive or T-tubing) can also affect this correlation Razi *et al.* (2012). Another way of estimation PaCO₂ is transcutaneous PCO₂ (PtcCO₂) analysis Eberhard (2007). PtcCO₂ estimates the PaCO₂ by measuring the CO₂ diffusion through body tissue and skin using a sensor on the skin surface. The PtcCO₂ can be better than the PetCO₂ in non-intubated intensive care unit (ICU) patients. In patients with acute respiratory failure Lermuzeaux *et al.* (2016), the PtcCO₂ showed a high correlation with PaCO₂ ($r = 0.97$) which was higher than the PetCO₂ - PaCO₂ correlation ($r = 0.62$). Unfortunately, transcutaneous techniques are not widely used because the sensors should be calibrated frequently, and they are only reliable in neonates who have a thin skin.

All studies mentioned above were done using a linear regression method. Many comparative studies proved that the machine learning techniques, where the MLP is part of, can provide better results than linear regression in the medical field Bhaduri *et al.* (2009); Martinez-Coll & Nguyen (2001); Shan & Jie (2011); Zernikow *et al.* (1998), financial field Naeini *et al.* (2010) and environmental field Adel Zaqoot *et al.* (2010); Choubin *et al.* (2016). Moreover, the MLP is proven to be very efficient in solving nonlinear problems, finding nonlinear relationships and filtering irrelevant inputs and noises Haykin (1998). Because some serious fluctuations can be observed in ABG in ICU Sood *et al.* (2010), it is crucial to be able to reliably estimate the PaCO₂ for critically ill patients to have a continuous monitoring and reduce the number of punctures.

Our primary aim was to estimate the PaCO₂ using non-invasive continuous measurements and the last ABG. Our hypothesis is that a new approach using an MLP neural network, non-invasive parameters from the ventilator and pulse oximetry can lead to a more accurate PaCO₂ prediction than standard methods based on V_{min} and PetCO₂.

3.2 Materials and methods

This retrospective study was done on a high-resolution database that collects data from a 32-bed tertiary pediatric intensive care unit (PICU) at Sainte-Justine Hospital (Montreal, Canada) Brossier, El Taani, Sauthier, Roumeliotis, Emeriaud & Juvet (2018). We included all patients aged under 18 years old between May 2015 and October 2018 mechanically ventilated. Patients had to have at least 2 ABGs measured to be included. There were no exclusion criteria. Data were automatically collected and stored in a secured research data warehouse. Automatic data collection methods were previously described Brossier *et al.* (2018). This project was approved by the institutional review board of the hospital (number 2019-2142). Because of the retrospective nature of the project, parental consent was waived.

3.2.1 Data

The data used in this work included :

1. **Arterial blood samples** : acquired at clinically indicated times, including the arterial blood CO₂ pressure (PaCO₂), the arterial blood O₂ pressure (PaO₂), the arterial oxygen saturation (SO₂), the arterial bicarbonate (HCO₃) and the lactic acid.
2. **Monitors and ventilators data** including the respiratory rate (RR), the inspired oxygen fraction (FiO₂), the positive inspiratory pressure over the positive end expiratory pressure (PIP), the mean airway pressure (MAwP), the expiratory tidal volume (V_{te}), the minute volume (V_{min}) and the transcutaneous oxygen saturation (SpO₂). These variables were collected every 30 seconds for all patients monitored in the ICU. These data were shown to be associated with PaCO₂ in the literature.

3. **Electronic medical records** : we used the demographic data such as gender, weight and age.

3.2.2 Datasets preparation

Since the blood gases are acquired at variable time interval and the other parameters are acquired every 30 seconds, it was not possible to match exactly every single set of parameters from the ventilator with an ABG. However, ABG timestamps include a small imprecision due to the time of the specimen to be sent to the analyzer, for example. To have more robust and representative values, we computed the median and the range of variation of the ventilator's parameters within a 15 min range before every blood gas. To palliate for the irregular ABGs frequency (for example, patients can have ABG measured every two hours or every 12 hours, depending on the patient's condition), we did four categories of blood gases : blood gases separated by a maximum of two hours, six hours, twelve hours and twenty-four hours. These categories were chosen to reflect clinical practice of ABG monitoring. If a patient had three or more blood gases in a time frame (2, 6, 12 or 24 hours), only the two with the largest time gap were considered. Because patients had a broad number of ABG during their stay that is associated with their severity, we split the train and test groups at patient level (a patient can participate in training or testing set but not in both) with stratification on the number of ABG divided into five quintiles categories : [0 - 20%[; [20% - 40%[; [40% - 60%[; [60% - 80%[and [80% - 100%]. In each subset, we randomly took 70% for the training set, 15% for the validation set and 15% for the testing set. All the inputs were normalized between 0 and 1 before being fed to the network and the output was rescaled to the normal range to be compared with the real values. To normalize the testing set, we used the same normalization applied in the training set.

Because many are using the PetCO₂ to approximate the PaCO₂, we added this parameter to our inputs to test the effect of this addition, when the data was available. Only patients with at least two PetCO₂ and PaCO₂ within 15 minutes or less from each other were considered (to be able

to use one to predict the other). These data were also divided in four categories (2h, 6h, 12h and 24h).

3.2.3 Multi-Layer Perceptron

The MLP Gardner & Dorling (1998) is one of the most common artificial neural network architectures. It is made of three main parts : input layer, hidden layer(s) and the output layer. It is a feedforward artificial neural network meaning that the connections between neurons do not form a cycle (each layer communicates exclusively with the next layer). To implement the MLP, we used the scikit-learn library (version 0.20.2 Pedregosa, Varoquaux, Gramfort, Michel, Thirion, Grisel, Blondel, Prettenhofer, Weiss, Dubourg, Vanderplas, Passos, Cournapeau, Brucher, Perrot & Duchesnay (2011)) in Python. The activation function used in our work was the rectified linear unit function (ReLU) because it is more robust to the vanishing gradient problem and more efficient than other activation functions such as logistic function or hyperbolic function. Weights were initialized using the Xavier method Glorot & Bengio (2010) :

$$W \sim U\left[-\frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}; \frac{\sqrt{6}}{\sqrt{n_j + n_{j+1}}}\right]$$

where W represents the weight, $U[-a; a]$ is the uniform distribution in the interval $(-a, a)$, n_j and n_{j+1} are the numbers of input and output nodes (of the layer that we're initializing its weights). It was shown that this initialization technique brings faster convergence while training the network.

The training of the network is done by using the backpropagation technique Rumelhart *et al.* (1988) in which the weights are updated in successive iterations in a way to minimize the loss of the output layer. The optimizer Adam is used in this process Kingma & Ba (2014). We used the mean squared error (MSE) as the loss function. As the training process may suffer from overfitting when the model learns the details specific to the training data that have a negative impact on new data, a regularization term was added to improve the performance of the network. We're using the L2 regularization (which is the sum of all the features weights) because, unlike

the L1 regularization, L2 can learn complex data patterns and is better when the output is a function of all the inputs.

To optimize the hyper parameters, we trained and validated networks with one, two and three hidden layers with 1 to 30 neurons in each layer to determine the best model. In all cases, networks with two hidden layers gave best results. As for the number of neurons per layer, it differs from a case to another.

To find the best combination of inputs, we tried all the possible combinations out of our available inputs : : calibration $PaCO_2$ (acquired at a previous time) , pH, oxygen arterial partial pressure, bicarbonate, lactic acid, respiratory rate, inspiratory fraction of oxygen, expired tidal volume, positive end-expiratory pressure, positive inspiratory pressure, dynamic compliance (tidal volume divided by positive inspiratory pressure), mean airway pressures, minute ventilation, time gaps and all the variable range of variation. For each combination, we did a grid search up to three hidden layers and 30 neurons per layer. The best combination with the highest percentage of estimations with an error less than 5 mmHg was considered the best.

To make our results more consistent, we used the leave one out cross validation ; each dataset was divided into five subsets also called folds, the model was trained on four of them and tested on the fifth that was left out. This process was repeated five times and the results were averaged to get the final results.

3.2.4 Traditional Techniques

The most used technique to approximate the $PaCO_2$ is the $PetCO_2$ approximation and recently the corrected $PetCO_2$ approximation Baudin *et al.* (2016). Clinicians may use the $PetCO_2$ as it is or they correct it using the following equation :

$$PaCO_2 = -1.525 + 0.876 \times PetCO_2 + 0.442 \times MAwP + 0.117 \times FiO_2 \quad (3.1)$$

All the PetCO₂ measurements that were associated with a PaCO₂ measurement were used to test this approach.

Another approach to approximate the PaCO₂ is by using the Vmin Wexler & Lok (1981) :

$$PaCO_{2t+1} = \frac{PaCO_{2t} \times Vmin_t}{Vmin_{t+1}} \quad (3.2)$$

3.2.5 Statistical Analyses

To validate our model and test its effectiveness, we measured the percentage of estimation with a PaCO₂ error smaller than ± 5 mmHg as the best accuracy and ± 10 mmHg as a moderate accuracy. These thresholds were clinically chosen after a survey among 20 intensivists at Sainte-Justine Hospital Center (Montreal, Canada) and based on the Clinical Laboratory Improvement Amendments recommendations Medicare (1992). This percentage will be considered as our main judge to assess the performance of the work. To test the efficiency of our models we will also use :

1. **Median Absolute Error MedAE** : this metric is used alongside the previous two to obtain a scaled error that is affected less than the mean squared error by the outliers.
2. **Pearson correlation coefficient.**
3. **Intraclass correlation coefficient (ICC)** In this study, the ICC (3,1) type is used since the rater, the MLP model, is a static model and the reliability is calculated from a single measurement Koo & Li (2016); Shrout & Fleiss (1979).

Finally, the **Bland Altman** plot Bland & Altman (1986) was used to visualize the agreement between the estimation and the real values and calculate the limits of agreement; Using the coefficient of determination alone cannot be used to assess agreement between two methods Bland & Altman (1986).

3.3 Results

We included 527 patients who had 17,329 ABGs. The female ratio was 45% (n=238). All patients had at least one episode of mechanical ventilation. The mortality of the patients participating in this study was 8% (n=42). Median age was 6.7 months (IQR 1 – 60) and median weight was 9.5 kg (IQR 4.1 – 27.1). Table 3.1 shows the characteristics of the patients (median and interquartile range) in the four categories 2h, 6h, 12h and 24h.

As shown in the table, the statistics of the ventilation parameters and the blood gases were almost the same among the categories. But there was a small difference in the demographic data.

Table 3.1 Patients characteristics - Median and interquartile range

Parameters	Predictive intervals between arterial blood gases			
	2h	6h	12h	24h
Patients (n)	295	478	460	425
Observations (n)	2339	6544	8499	6214
<i>Demographic data</i>				
Age (months)	7.6 (1.2-78.9)	7.6 (1.4-67.2)	6.7 (1.3-61.8)	6.6 (1.3-62.1)
Weight (kg)	8.9 (4.35-23.0)	7.82 (4.35-22.2)	7.3 (4.28-20.0)	7.4 (4.28-20.0)
Female (%)	45.4	45.2	45.0	44.6
<i>Ventilation parameters</i>				
RR (/min)	30 (20-38)	32 (23-38)	32 (24-38)	32 (24-38)
FiO ₂ (%)	53 (39-78)	50 (39-71)	50 (36-67)	49 (35-64)
PIP (cmH ₂ O)	23 (19-27)	22 (18-26)	21 (17-26)	21 (17-25)
MAwP (cmH ₂ O)	13 (10-15)	12 (10-15)	12 (9-14)	12 (9-14)
Vte (ml/kg)	6 (5-8)	6 (5-7)	6 (5-7)	6 (5-7)
Vmin (L/min)	1.5 (0.8-4.2)	1.6 (0.9-3.7)	1.5 (0.8-3.5)	1.5 (0.8-3.6)

Table 3.1 Patients characteristics - Median and interquartile range (continued)

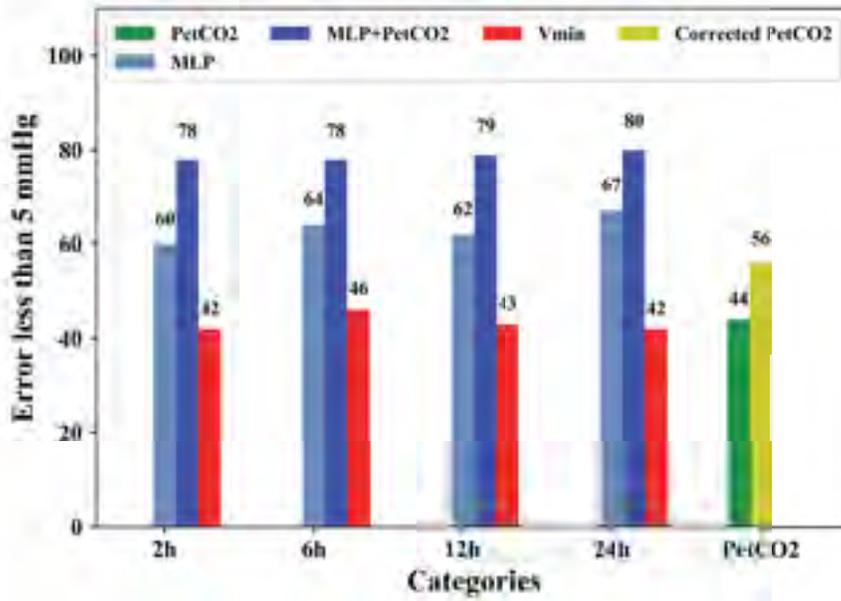
Parameters	Predictive intervals between arterial blood gases			
	2h	6h	12h	24h
<i>Continuous monitoring</i>				
SpO ₂ (%)	98 (94-100)	98 (95-100)	98 (95-100)	98 (95-100)
<i>Biological data</i>				
Blood gas				
PaCO ₂ (mmHg)	43 (37-50)	42 (37-49)	43 (38-49)	43 (38-49)
PaO ₂ (mmHg)	96 (67-163)	96 (69-142)	95 (69-138)	95 (70-136)
SaO ₂ (%)	98 (93-99)	98 (94-99)	98 (94-99)	98 (95-99)
HCO ₃ (mmol/L)	26 (22-29)	26 (23-29)	26 (23-30)	27 (23-30)
Lactic acid (mmol/L)	1.6 (1.0-2.9)	1.3 (0.9-2.1)	1.2 (0.87-1.9)	1.2 (0.8-1.8)
pH	7.39 (7.31-7.46)	7.4 (7.34-7.46)	7.41 (7.34-7.46)	7.41 (7.36-7.45)
<i>Clinical outcome</i>				
Mortality rate (%)	12.2	8.6	8.5	8.7
<p>RR : respiratory rate, FiO₂ : inspired oxygen fraction, PIP : Positive inspiratory pressure, MAwP : mean airway pressure, Vte : expiratory tidal volume, Vmin : minute volume, SpO₂ : transcutaneous oxygen saturation, PaCO₂ : arterial blood CO₂ pressure, PaO₂ : arterial blood O₂ pressure, SaO₂ : arterial oxygen saturation, HCO₃ : arterial bicarbonates</p>				
<p>It should be noted that the same patient can be present in many categories ; therefore the categories are not demographically independent.</p>				

After splitting the datasets into training (around 70%), validation (around 15%) and testing (around 15%), the testing sets had between 14.9% and 16% of the patients with 15.5% and 15.7% of the observations, and the PetCO₂ testing sets had between 15.5% and 16.8% of the patients with 14.8% to 16% of the observations.

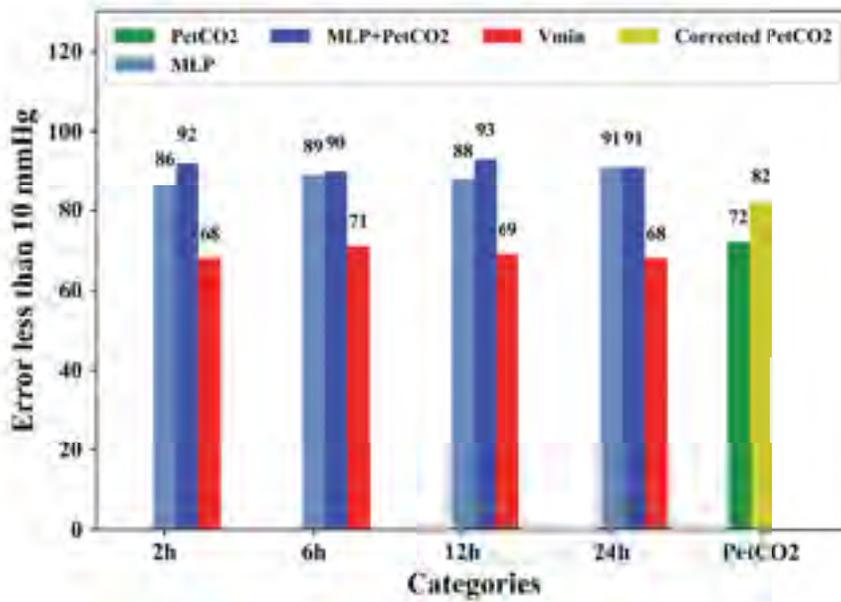
3.3.1 Variables selection :

We found that adding the range of variation to the inputs for some combinations of inputs increased the accuracy of the estimations and for others decreased the accuracy. In both cases the changes were very minimal (smaller than 0.5% in the percentage of estimations with an error less than 5 mmHg and 10 mmHg). So, the range of variation was excluded from the inputs. Also, adding the pH, PaO₂, HCO₃ and the lactate or any combination of these parameters to the inputs had a negligible increase or decrease in the percentage of estimation with an error less than 5 mmHg and less than 10 mmHg (less than 0.2% in all 4 categories). Only PaCO₂ from the blood gas data was included in the inputs. The best combination was : PaCO₂, RR, FiO₂, PIP, MAwP, Vmin and Time Gap.

In table 3.2 we present the median absolute error with the interquartile ranges, the R², the ICC with its 95% confidence intervals and the size of the testing sets for all the techniques that we tried in this study. The accuracy of the predictions represented by the percentage of estimations with an error less than 5 mmHg and 10 mmHg is shown in figure 3.1. Our MLP model outperformed all other techniques with accuracy *leq* 5 mmHg of 60%, 64%, 62% and 67% throughout the categories while the Vmin approximation gave an accuracy of 42%, 46%, 43% and 42% and the PetCO₂ approximation an accuracy of 56%. The median absolute error was the lowest with the MLP models and, overall, the Pearson correlation coefficient for the MLP models was higher than in the other techniques. Including the PetCO₂ to the inputs in our approach improved the accuracy of the estimations by 13% to 18% (percentages of errors less than 5 mmHg. The Pearson correlation coefficient and the intra-class correlation coefficient also improved. This improvement should be investigated further to confirm that the PetCO₂ addition caused it since the number of data and patients were very few compared to the other datasets. We should mention that the “numbers of ABGs used for testing” for the MLP models in the table 3.2 are around 15% of the whole dataset that was split into training and testing.



A. Percentage of errors ≤ 5 mmHg



B. Percentage of errors ≤ 10 mmHg

Figure 3.1 Overall accuracy of the predictive models

Table 3.2 Prediction validation

Time gap with calibration ABG	Model	MedAE (IQR) mmHg	R2	ICC (95% CI)	Number of ABGs used for testing	Number of patients
0h	PetCO ₂	4.3 (1.9-8.2)	0.32	0.56 (0.53-0.59)	2346	441
2h	Vmin	6.3 (2.7-12.4)	0.05	0.09 (0.05-0.13)	2530	299
	MLP	4.0 (2.0-7.6)	0.37	0.54 (0.47-0.61)	396	45
	MLP + PetCO ₂	3.4 (2.0-5.0)	0.59	0.75 (0.57-0.93)	25	11
6h	Vmin	5.7 (2.5-11.1)	0.03	0.06 (0.04-0.08)	6521	478
	MLP	3.7 (1.8-6.7)	0.44	0.58 (0.54-0.62)	1024	76
	MLP + PetCO ₂	2.9 (1.7-5.2)	0.53	0.68 (0.54-0.82)	59	21
12h	Vmin	5.9 (2.7-11.7)	0.02	0.05 (0.03-0.07)	8472	460
	MLP	3.8 (1.8-6.9)	0.31	0.5 (0.46-0.54)	1313	69
	MLP + PetCO ₂	3.0 (1.5-5.0)	0.64	0.75 (0.64-0.86)	61	24
24h	Vmin	6.2 (2.8-12.2)	0.01	0.02 (0.0-0.04)	6704	460
	MLP	3.5 (1.6-6.2)	0.4	0.57 (0.53-0.61)	1053	68
	MLP + PetCO ₂	3.0 (1.6-5.0)	0.48	0.66 (0.52-0.8)	64	21
ABG : arterial blood gases, Vmin : minute volume ventilation, MLP : Multilayer perceptron, ICC : Intraclass correlation						

To study the effect of the patient's clinical situation on the accuracy of the estimation, we separated the patients into two groups depending on the patient's PaCO₂ situation : if the difference between the PaCO₂ at the time of calibration and the PaCO₂ at the time of estimation (both are

measured values and not estimations) was smaller than 5 mmHg then PaCO₂ was considered as unchanged PaCO₂, and if the difference was larger than 5 mmHg the PaCO₂ was considered as changed PaCO₂. Table 3.3 shows the accuracy of the models according to the patients' situation.

Table 3.3 Accuracy (≤ 5 mmHg) robustness comparison when significant changes occur between calibration ABG and predicted PaCO₂

Time gap with calibration ABG	Model	Accuracy if PaCO ₂ was unchanged	Accuracy if PaCO ₂ changed
2h	Vmin	69%	14%
	MLP	77%	42%
	MLP + PetCO ₂	86%	61%
6h	Vmin	63%	24%
	MLP	85%	35%
	MLP + PetCO ₂	91%	47%
12h	Vmin	57%	31%
	MLP	82%	43%
	MLP + PetCO ₂	87%	63%
24h	Vmin	51%	29%
	MLP	86%	39%
	MLP + PetCO ₂	90%	67%
Vmin : minute volume ventilation, MLP : Multilayer perceptron			

The patients' situation has an impact on the accuracy of the estimation. For all approaches, the accuracy if the PaCO₂ was not changed was way higher than the accuracy if PaCO₂ was changed. In both situations, the MLP models outperformed the Vmin approximation in all four

categories. This table also showed that the MLP + PetCO₂ had the best accuracy. However, more tests should be done in this regard.

The data were divided in three subgroups depending on the values of the PaCO₂ : 1) normocapnic patients with a PaCO₂ in the normal range [35 mmHg; 45 mmHg], 2) hypocapnic patients with PaCO₂ < 35 mmHg and 3) hypercapnic patients with PaCO₂ > 45 mmHg (based on the measured values and not the estimated ones). The results are shown in table 3.4.

Table 3.4 Accuracy (< 5 mmHg) comparison in hypo-, normo- and hypercapnia

Time gap with calibration ABG	Model	< 35 mmHg	[35-45] mmHg	> 45 mmHg	Global Accuracy
<i>0h</i>	PetCO ₂	50%	65%	43%	56%
<i>2h</i>	Vmin	50%	48%	35%	42%
	MLP	31%	78%	51%	60%
	MLP + PetCO ₂	83%	78%	58%	78%
<i>6h</i>	Vmin	48%	51%	38%	46%
	MLP	38%	79%	51%	64%
	MLP + PetCO ₂	73%	89%	59%	78%
<i>12h</i>	Vmin	43%	50%	36%	43%
	MLP	24%	77%	54%	62%
	MLP + PetCO ₂	60%	93%	73%	79%
<i>24h</i>	Vmin	46%	46%	37%	42%
	MLP	25%	79%	62%	67%
	MLP + PetCO ₂	60%	91%	73%	80%
<p>ABG : arterial blood gases, Vmin : minute volume ventilation, MLP : Multilayer perceptron</p>					

In all techniques, the best accuracy was obtained in the normocapnia, then in the hypercapnia and last in the hypocapnia cases. Some fluctuations can be found in the MLP+PetCO₂ which is probably due to the small size of the testing sets for some subgroups folds (fewer than 5 observations). Overall and by excluding the MLP+PetCO₂ results, the MLP outperformed the traditional techniques for the normo- and hypercapnic patients while the PetCO₂ approximation was best for the hypocapnic patients.

In figure II we present the Bland Altman plots. We chose to plot the results from the 6h category for our MLP approach with and without PetCO₂ from one of the 5 folds ; the third plot represents the PetCO₂ approximation (eq 2) and the last plot represents the results of the Vmin approximation (eq 1) in the 6h category.

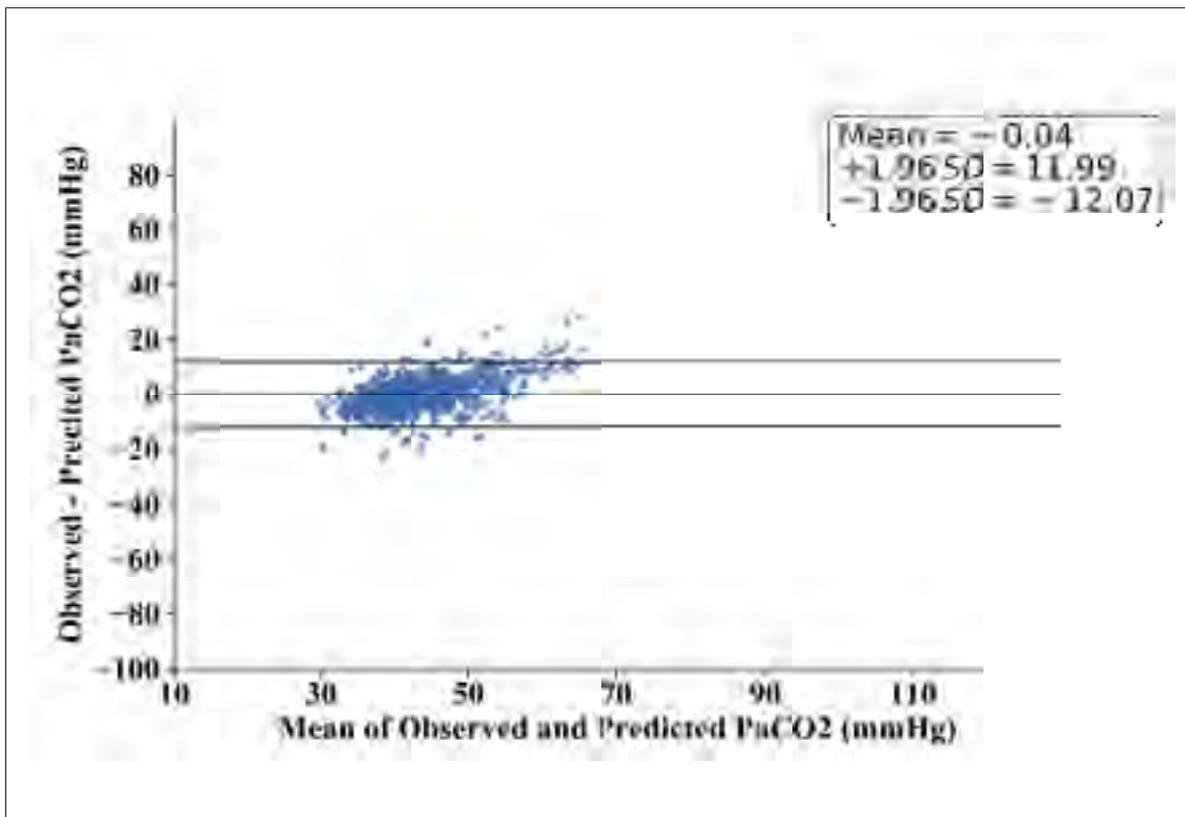


Figure 3.2 Bland Altman plot of the MLP technique in the 6h category

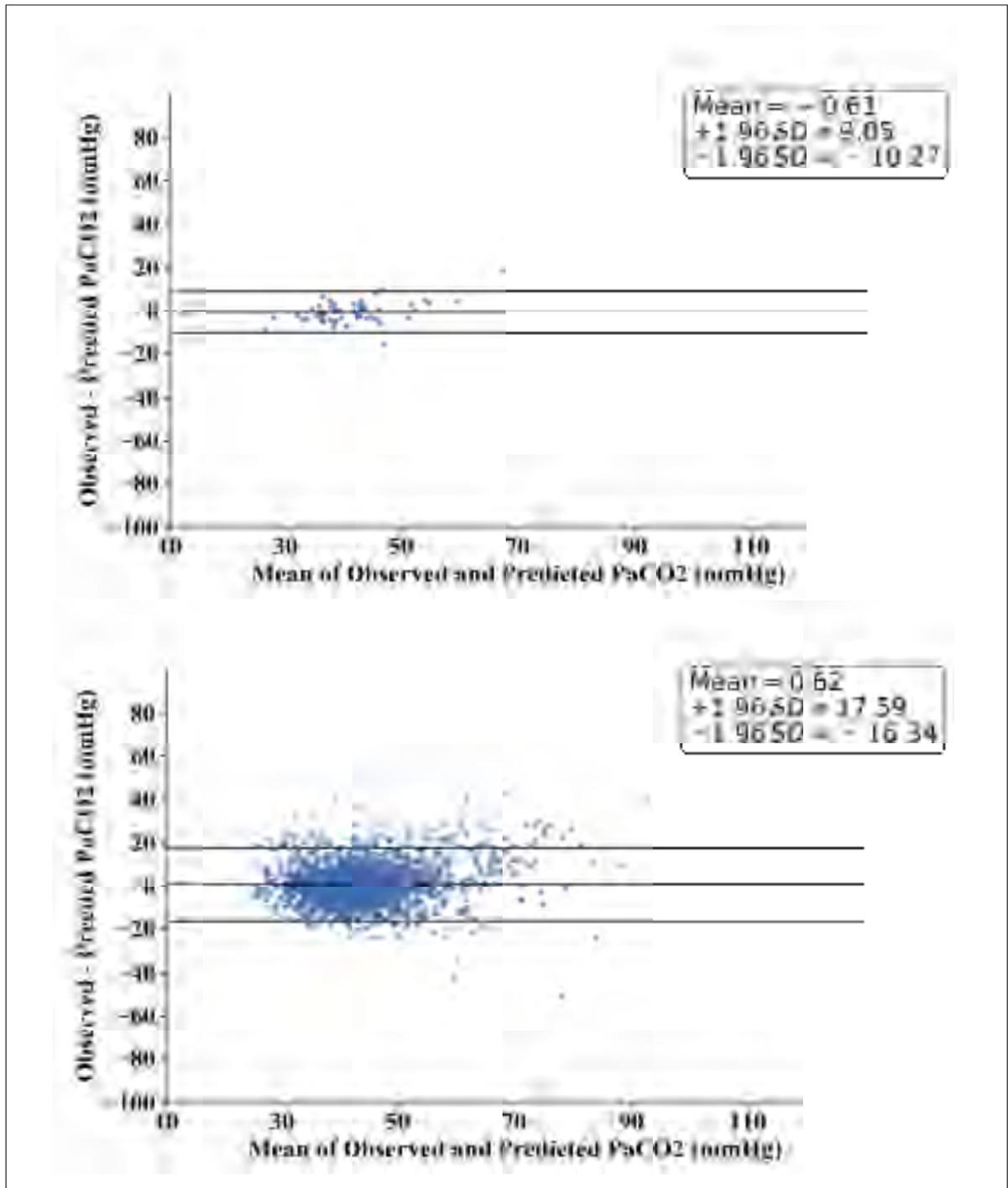


Figure 3.3 Bland Altman plots of the MLP +PetCO₂ technique in the 6h category (top) and the PetCO₂ approximation - equation 3.1 (bottom)

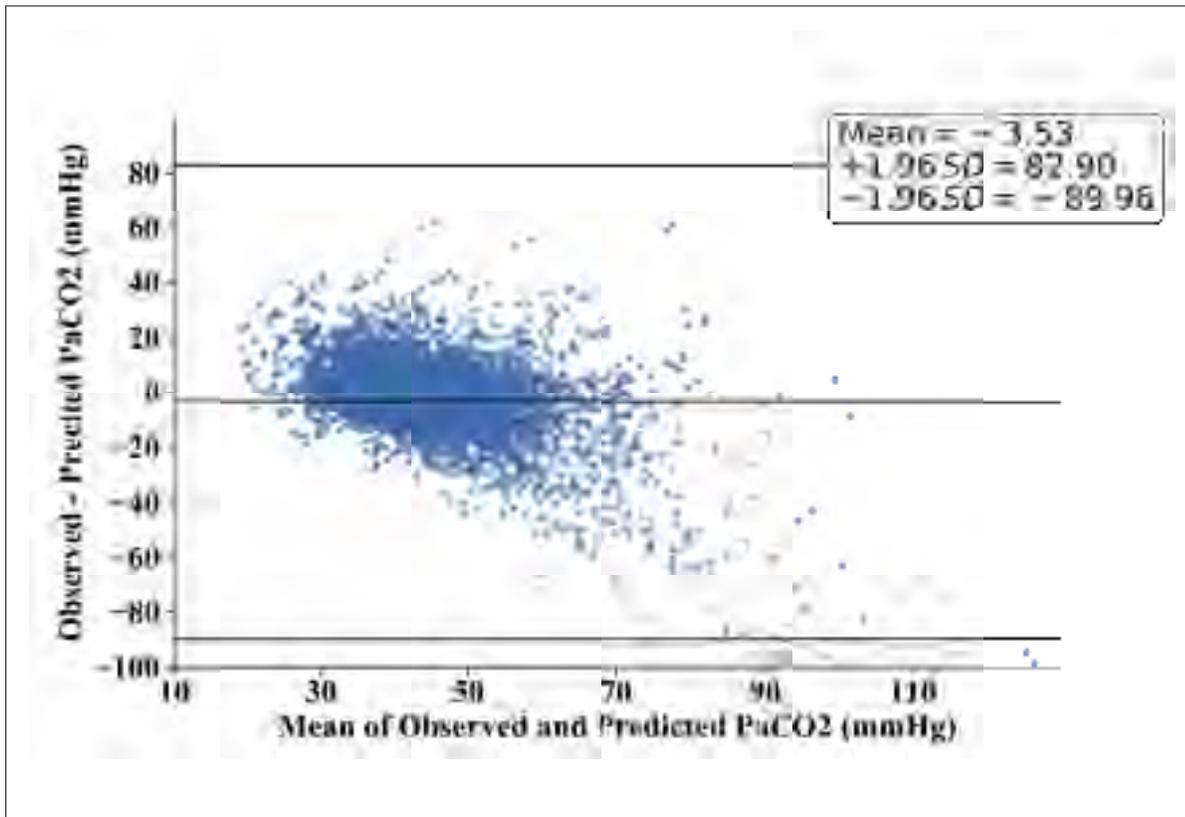


Figure 3.4 Bland Altman plot of the V_{min} approximation in the 6h category - equation 3.2

The characteristics of the Bland Altman (biases, limits of agreement and their confidence intervals) in the categories for all techniques are shown in table 3.5. As we did in the results, we computed the Bland Altman characteristics in the 5 folds in each category and then we averaged the results that are shown in the table. The MLP technique had the most consistent Bland Altman plots and the confidence intervals were the smallest compared to the other techniques. The MLP + $PetCO_2$ may have limits of agreements but its confidence intervals were way bigger than the confidence intervals of the MLP. The $PetCO_2$ approximation was second best, as for the V_{min} approximation, it had the worst characteristics and the biggest confidence intervals.

Table 3.5 Bland Altman Characteristics
Bias, upper and lower 95% limits of agreement and their 95% confidence intervals

		Bias (mmHg)	Upper limit (mmHg)	Lower limit (mmHg)
<i>PetCO₂</i>	0h	0.45 ±1.12	18.25 ±1.92	-17.35 ±1.92
<i>Vmin</i>	<i>2h</i>	-3.62 ±4.89	43.55 ±8.39	-50.78 ±8.39
	<i>6h</i>	-3.71 ±4.52	68.18 ±7.74	-75.59 ±7.74
	<i>12h</i>	-4.23 ±4.16	70.96 ±7.12	-79.43 ±7.12
	<i>24h</i>	-2.51 ±2.69	39.51 ±4.61	-44.53 ±4.61
<i>MLP</i>	<i>2h</i>	0.19 ±1.61	16.20 ±2.76	-15.82 ±2.76
	<i>6h</i>	0.28 ±0.85	13.90 ±1.46	-13.35 ±1.46
	<i>12h</i>	0.34 ±0.77	14.27±1.31	-13.58 ±1.31
	<i>24h</i>	0.01 ±0.76	12.40 ±1.31	-12.38 ±1.31
<i>MLP + PetCO₂</i>	<i>2h</i>	-1.45 ±4.83	10.51 ±8.82	-13.42 ±8.82
	<i>6h</i>	0.85 ±3.10	12.82 ±5.44	-11.12 ±5.44
	<i>12h</i>	1.33 ±2.70	11.87 ±4.73	-9.21 ±4.73
	<i>24h</i>	0.02 ±2.64	10.53 ±4.62	-10.49 ±4.62
Vmin : minute volume ventilation, MLP : Multilayer perceptron				

3.4 Discussion

In this retrospective study using data collected prospectively, we showed that the PaCO₂ can be estimated using artificial neural networks. We also showed that this method performed better than the already existing methods, such as minute ventilation based and end-tidal CO₂-based methods. Overall patients with all PaCO₂ ranges, the predictions were between 60% to 67% with an error less than 5 mmHg and 86% to 91% with an error less than 10 mmHg.

Managing patients in the ICU requires blood gas analysis which gives information about the gas exchange situation and the acidity of the blood Organization *et al.* (2010); Sood *et al.* (2010). The most accurate way to measure the ventilation state is to intermittently draw arterial blood gas samples. Our approach tends to provide a valid estimation with less invasive measurements. The

main advantage of this approach is to rely on already commonly used devices in intensive care where most of the arterial blood gases are performed, such as the monitors and the ventilator. Also, it can be applicable in all situations. This approach is more critical for children who have smaller arteries, resulting in more difficult ABG access. Once finalized, the model can be implemented in clinical decision support systems with an option to include the latest blood gas acquisition and the estimations will be shown continuously.

Variable analysis showed that adding the age or the weight or both to the inputs didn't have a significant effect on the estimation, except that tidal volumes were normalized with weight (ml/kg) as it is standard practice in pediatric critical care. Others found similar results Panigrahi *et al.* (2018), and one can hypothesize that despite physiological changes related to growths, such machine learning algorithms can overcome this. We also found that the models with the MAWP and the FiO_2 yield better results than the ones without these two parameters. These variables are key in the estimation of the severity of the lung injury Khemani, Smith, Zimmerman, Erickson, Group *et al.* (2015). Thus, we hypothesized that these variables are used by the network to estimate how severe is the lung disease and how they are able to eliminate carbon dioxide.

We also found that our network performed better to predict normal results ($PaCO_2$ in the normal range [35 mmHg, 45 mmHg]) with 77% to 79% of estimations with an error less than 5 mmHg. As the normal category was the most frequent one (44% to 49% of all values), such results were expected. However, these results were better than currently available methods to estimate $PaCO_2$; the V_{min} approximation had between 46% and 51% of estimations with an error less than 5 mmHg all the categories while the $PetCO_2$ approximation had 65%. Moreover, our model also performed well for hypercapnic patients ($PaCO_2 > 45$ mmHg) with 51% to 62% of the estimations with an error less than 5 mmHg. These results were also better than existing models (35% to 38% for the V_{min} approximation and 43% for the $PetCO_2$ estimation). For hypocapnic patients ($PaCO_2 < 35$ mmHg), our models had a lower performance (24% to 38% of the estimations with an error less than 5 mmHg) compared to existing models (43% to 50% for the V_{min} and 50% for the $PetCO_2$). This result could be due to the unbalanced nature of our

data specially for the hypocapnic cases where fewer data were found, in fact, clinicians in ICU intervene immediately when the patient is hypocapnic and they react less for other cases.

Some clinicians may estimate the PaCO₂ by measuring the gap between the PaCO₂ and PetCO₂ at a certain time point and keep the difference constant. However, the error between PetCO₂ and PaCO₂ is not constant over time as shown in Figure 2 (panel C) and by other studies in the field Baudin *et al.* (2016); Rentola, Hästbacka, Heinonen, Rosenberg, Häggblom & Skrifvars (2018).

One strength of our model is that it does not require the end-tidal CO₂ variable to be included. Such variable requires additional measuring equipment and may not always be available. However, when added to the model, the accuracy of the estimations improved by 13% to 18% (percentage of estimations with an error less than 5 mmHg) and this improvement was in all 4 categories and for all patients with a changed or unchanged PaCO₂ situation. However, these results were obtained with a small number of patients and a small number of observations that were unbalanced changed-unchanged patients wise ; therefore, more tests should be done with more data to conclude. A close study to ours also showed that machine learning (Gaussian process) can be useful in the PaCO₂ estimation based on noninvasive parameters from the ventilator's settings along with the PetCO₂ Khemani, Celikkaya, Shelton, Kale, Ross, Wetzel & Newth (2014). The authors found that 80% of the predictions had an error less than 6 mmHg.

To better interpret the robustness of the model in case of changed patients, we separated the cases between changed and unchanged PaCO₂ results (Table 3.3). As expected, the performance of our model was lower when the PaCO₂ changed than when it was constant. However, the estimation was better using our model than the already available ones. We interpret this as how complex is dynamic physiology. In addition, we observed a better estimation for changed patients when the time gap between the calibration ABG and the predicted ABG was short. Despite that, the difference is relatively small.

Our study has limitations. First, our model needs to be calibrated on a previous ABG, making our technique less applicable outside of a hospital setting and requiring an ABG test that we

are trying to reduce its numbers. Because the sickest patients already may have an arterial catheter for other reasons (for example continuous blood pressure monitoring), the impact of this limitation could be minimized for those patients. Our method still provides the advantage of continuous estimation of the PaCO₂ for hospitalized patients, which cannot be provided by ABG tests. Second, despite having better results than the already in use PaCO₂ estimation techniques, the accuracy of our approach should be improved further before being completely adopted. Third, the performance of the model is lower in case of extreme severity. In consequence, our tool should not be used without ABG confirmation in case of suspected sudden clinical change or extreme severity. Again, as our model requires previous ABG to calibrate, clinicians should also validate the estimation of the model before relying on it. Fourth, we should also mention that the hemodynamic state was not continuously assessed in this study, cardiac output may explain some of the error between the PetCO₂ and the PaCO₂. Finally, although this study has the potential to decrease the number of required ABG in patients who require mechanical ventilation, other indications of drawing ABG subsist, such as pH or arterial partial pressure evaluation.

The influence of the PetCO₂ on the estimation should be retested. Acquiring more PetCO₂ data from many unchanged and changed patients having many PaCO₂ in the three subgroups are needed to verify whether the result found above can be generalized. When verified, adding a PetCO₂ option to the tool can be very helpful in the estimations.

Moreover, a more complete and accurate tool can be built ; during our research for the best inputs combination, we found many different best combinations, each for a subgroup (< 35 mmHg, [35 mmHg, 45 mmHg], > 45 mmHg for patients with a stable and unstable PaCO₂ situation). Developing a model that can classify each observation in one of the six classes can improve the performance of the estimation. Acquiring more data is needed to develop such a classifier. Once classified, the appropriate class model can be used to calculate the estimation. By applying the right input combination for each subgroup, we increase the estimations with an error less than 5 mmHg from 77% to 96% for patients with unchanged situation and from 42% to 65% for patients with a changed situation, so overall from 60% to 82%.

Including the dead space to tidal volume ratio (V_D/V_T) or the respiratory index SpO_2/FiO_2 can also be an option to improve the accuracy of the estimation; these 2 factors are proven to have a great influence on the estimation Panigrahi *et al.* (2018); Wexler & Lok (1981).

More data are needed to test and apply these additions which won't be a problem since the databases are growing bigger and bigger each day.

3.5 Conclusion

We proposed in this work a reliable and immediate model capable of estimating the $PaCO_2$ using only noninvasive ventilator parameters calibrated on a previous arterial blood gas test. The model showed a good accuracy for the normally ventilated patients and tended to decrease for hypercapnic and hypocapnic patients. However, our multilayer perceptron was shown to be better than existing model to estimate $PaCO_2$ such as the V_{min} and the $PetCO_2$ approximations. Implementing this model in ICUs could allow clinicians to continuously monitor the $PaCO_2$ and to reduce the number of ABG tests leading to a better treatment of patients and lesser harm dealt.

3.6 Acknowledgement

This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC), in part by the Institut de Valorisation des données de l'Université de Montréal (IVADO) and in part by the Fonds de la recherche en sante du Quebec (FRQS). (corresponding author : Jihad El Tannoury)

CONCLUSION

L'objectif principal des travaux présentés dans ce mémoire est de trouver une façon fiable noninvasive pour estimer les pressions artérielles du CO₂. Ces pressions sont principalement utilisées dans l'analyse de la respiration et de la situation acido-basique du sang et sur lesquelles les cliniciens se basent pour ajuster les paramètres de la ventilation mécanique. La procédure standard pour obtenir la PaCO₂ consiste à faire un prélèvement du sang artériel. Cette technique est douloureuse et peut endommager les vaisseaux et les nerfs dans l'entourage et parfois causer des infections ; d'où l'importance de trouver une alternative noninvasive.

En pratique, plusieurs alternatives sont utilisés pour estimer les PaCO₂ ; 1) la PetCO₂ et sa forme corrigée sont utilisées dans l'estimation, mais ces techniques ne peuvent pas être toujours applicables (cela dépend des conditions médicales du patient) et pour obtenir la PetCO₂ il faut ajouter des machines supplémentaires au traitement, 2) la pression transcutanée du CO₂ obtenue par des capteurs sur la peau, mais ces capteurs doivent être calibrés fréquemment et sont uniquement fiables chez les nouveau-nés dotés d'une peau mince, 3) la PexCO₂ dans les circuits extracorporels qui sont des cas extrêmement rares, et 4) le rapport de proportionnalité basé sur le volume minute qui peut être un alternatif efficace mais strictement dans des cas spécifiques.

La majorité des techniques en pratique utilise la régression linéaire simple ou multiple pour faire les estimations. Rares sont les études qui utilisent les algorithmes de l'intelligence artificielle. Toutefois, ces algorithmes ont montré un grand succès dans la prédiction et l'estimation des paramètres dans tous les domaines (finance, environnement, santé . . .).

Dans notre travail, on a utilisé la RR, la FiO₂, la PIP, la MAwP et le Vmin pour estimer la PaCO₂. Ces paramètres sont noninvasifs et surveillés d'une manière continue dans les PICUs. On a

utilisé le multilayer perceptron calibré par la PaCO₂ obtenue par un test ABG réalisé auparavant pour faire les estimations. D'autre part, on a fait l'estimation du PaCO₂ en utilisant la PetCO₂ dans sa forme corrigée dans un second lieu et utilisant le Vmin dans un troisième lieu.

Les résultats obtenus par ces trois approches montrent que l'approche MLP utilisant les paramètres noninvasifs a performé mieux que les deux autres approches ; le pourcentage des estimations avec une erreur moins que 5 mmHg était supérieur à ceux des deux autres techniques et les caractéristiques des plots Bland-Altman étaient meilleures et plus consistantes. Ajouter la PetCO₂ aux paramètres noninvasifs dans notre approche a amélioré davantage la performance de l'estimation, mais ce résultat est obtenu avec très peu de données et il vaut mieux répéter l'étude avec plus de données pour s'assurer de ce résultat.

Pour mieux étudier la performance de notre approche, on a divisé les patients sur deux groupes selon l'évolution de la PaCO₂ : patients avec PaCO₂ changée et inchangée ; puis chaque groupe était divisé en trois sous-groupes selon la situation du PaCO₂ : hypo-, normo- et hypercapnie. Les résultats montrent que la performance dans le groupe des PaCO₂ inchangées était meilleure et que dans chaque groupe, la performance est meilleure dans la normocapnie, puis l'hypercapnie et puis dans l'hypocapnie.

En conclusion, on peut dire qu'on a développé un modèle capable d'estimer la PaCO₂ d'une manière efficace et fiable dont la performance a dépassé celle des techniques utilisées aujourd'hui en pratique clinique.

RECOMMANDATIONS

Au cours de ce travail de recherche, on a pu formuler plusieurs observations et des recommandations pour ceux qui souhaitent continuer et améliorer le travail, on cite :

Une Classification

Dans ce projet, on a traité toutes les données d'une manière identique et on a utilisé un seul réseau pour faire les estimations. Toutefois au cours de la phase de sélection de la meilleure combinaison de paramètres (celle qui a donné le meilleur pourcentage d'estimations avec une erreur < 5 mmHg), on a trouvé que pour chaque groupe de patients : ayant une $\text{PaCO}_2 < 35$ mmHg, > 45 mmHg ou entre ces deux valeurs et ayant une situation changée ou non, on a trouvé que chacun de ces six groupes possède une meilleure combinaison de paramètres différents des autres groupes. Un travail de classification fait auparavant peut donc améliorer beaucoup les résultats. On n'a pas pu réaliser ce travail puisque les données n'étaient pas suffisantes et il fallait une plus grande base de données pour faire un classifieur efficace. Donc, faire beaucoup plus d'acquisitions dans ces six groupes et ensuite faire un modèle qui classifie les données dans un de ces groupes peut largement influencer les résultats. Si on applique la classification à notre travail, le pourcentage d'estimations avec une erreur < 5 mmHg augmente de 77% à 96% pour les patients avec une situation inchangée et de 42% à 65% pour ceux qui ont une situation changée et alors en total une amélioration de 60% à 82% pour la catégorie de 2h.

Un Ajout de paramètres

On a constaté qu'ajouter la PetCO_2 aux paramètres utilisés dans le réseau améliore aussi les estimations et ceci pour toutes les quatre catégories. Ce résultat a été trouvé sur peu de données et pour l'adopter, on suggère de retester l'influence de l'ajout du PetCO_2 sur les résultats en faisant

une acquisition plus grande de données $P_{et}CO_2$ pour des patients avec une situation changée et inchangée et ayant du P_aCO_2 dans les trois marges. De plus, il serait aussi bénéfique de faire la classification, expliquée juste en haut, de ces données $P_{et}CO_2$.

Une autre addition peut être considérée c'est l'addition du rapport espace mort sur le volume tidal (V_D/V_T) et l'indice de respiration défini par P_aO_2/FiO_2 . Panigrahi *et al.* (2018); Wexler & Lok (1981) ont montré que ces deux rapports ont une grande influence sur l'estimation du P_aCO_2 et les prendre en considération peut améliorer les résultats. Par les prendre en considération on veut dire distribuer les patients sur des groupes suivant des marges de ces deux rapports et non pas les inclure dans les paramètres utilisés par le réseau pour faire l'estimation. Par la suite, comparer les résultats de cette distribution à celle qu'on a adoptée dans nos travaux pour voir laquelle sera la plus efficace.

Un modèle complet

Notre vision future de ce travail est de créer un modèle complet qui utilise les paramètres disponibles pour faire l'estimation du P_aCO_2 . Au départ, une classification sera faite pour savoir quelle combinaison de paramètres à inclure, puis le modèle correspondant sera utilisé pour faire l'estimation. Si la $P_{et}CO_2$ est mesurée, les modèles qui gèrent les $P_{et}CO_2$ seront utilisés après avoir classé les données dans le groupe correspondant. Aussi, les méthodes traditionnelles seront utilisées pour les cas où elles ont performé mieux que notre approche : comme l'approximation V_{min} dans les cas d'hypocapnies.

ANNEXE I

RÉSULTATS DES 5 FOLDS POUR LES MODÈLES MLP

Table-A I-1 Accuracy (Accuracy (≤ 5 mmHg) robustness comparison when significant changes occur between calibration ABG and predicted PaCO₂)

Time gap with calibration ABG	Accuracy if PaCO ₂ was unchanged	Accuracy if PaCO ₂ changed	
2h -> Fold nb	1	77%	41%
	2	78%	48%
	3	77%	40%
	4	75%	38%
	5	78%	42%
Average	77%	42%	
6h -> Fold nb	1	86%	36%
	2	88%	30%
	3	83%	35%
	4	82%	37%
	5	87%	36%
Average	85%	35%	
12h -> Fold nb	1	86%	41%
	2	82%	43%
	3	78%	50%
	4	82%	41%
	5	80%	40%
Average	82%	43%	
24h -> Fold nb	1	82%	43%
	2	88%	36%
	3	88%	36%
	4	86%	41%
	5	87%	36%
Average	86%	39%	

Table-A I-2 Prediction validation

Time gap with calibration ABG	MedAE (IQR)	R2 (95% CI)	ICC (95% CI)	Number of ABGs used for testing	Number of patients
2h → Fold nb	1 4.3 (2.4-7.8)	0.56 (0.48-0.62)	0.51 (0.43-0.58)	386	48
	2 4.0 (2.0-7.2)	0.62 (0.55-0.67)	0.58 (0.51-0.64)	388	46
	3 3.8 (1.9-7.2)	0.59 (0.52-0.65)	0.51 (0.43-0.58)	398	45
	4 4.3 (2.0-8.1)	0.68 (0.63-0.73)	0.60 (0.53-0.66)	390	43
	5 3.6 (1.9-7.8)	0.60 (0.53-0.66)	0.51 (0.43-0.57)	419	44
Average	4.0 (2.0-7.6)	0.61 (0.54-0.67)	0.54 (0.47-0.6)	396	45
6h → Fold nb	1 3.7 (1.8-6.7)	0.65 (0.61-0.69)	0.57 (0.53-0.61)	1011	73
	2 3.9 (2.0-7.2)	0.54 (0.5-0.58)	0.49 (0.44-0.54)	1068	78
	3 3.7 (1.8-6.8)	0.69 (0.66-0.72)	0.60 (0.56-0.64)	1008	78
	4 3.7 (1.7-6.6)	0.69 (0.65-0.72)	0.63 (0.59-0.66)	996	76
	5 3.5 (1.8-6.3)	0.71 (0.68-0.74)	0.62 (0.58-0.65)	1035	77
Average	3.7 (1.8-6.7)	0.66 (0.62-0.69)	0.58 (0.54-0.62)	1024	76
12h → Fold nb	1 3.7 (1.8-6.7)	0.64 (0.61-0.67)	0.55 (0.51-0.58)	1331	66
	2 3.7 (1.8-6.8)	0.52 (0.47-0.55)	0.46 (0.42-0.51)	1286	70
	3 3.7 (1.8-6.5)	0.61 (0.57-0.64)	0.53 (0.49-0.56)	1303	70
	4 4.0 (1.8-7.2)	0.43 (0.39-0.48)	0.41 (0.36-0.45)	1327	66
	5 4.0 (1.9-7.4)	0.59 (0.55-0.62)	0.53 (0.49-0.57)	1316	71
Average	3.8 (1.8-6.9)	0.56 (0.52-0.59)	0.50 (0.45-0.53)	1313	69

Table-A I-2 Prediction validation (continued)

Time gap with calibration ABG	MedAE (IQR)	R2 (95% CI)	ICC (95% CI)	Number of ABGs used for testing	Number of patients
<i>24h -> Fold nb</i> 1	3.5 (1.6-6.4)	0.59 (0.55-0.62)	0.51 (0.46-0.55)	1102	68
2	3.3 (1.3-6.1)	0.62 (0.58-0.65)	0.57 (0.53-0.61)	1024	65
3	3.6 (1.7-6.2)	0.61 (0.57-0.65)	0.56 (0.52-0.6)	1015	68
4	3.5 (1.7-5.7)	0.69 (0.66-0.72)	0.63 (0.6-0.67)	1060	67
5	3.6 (1.7-6.5)	0.62 (0.58-0.65)	0.57 (0.52-0.61)	1063	71
<i>Average</i>	3.5 (1.6-6.2)	0.63 (0.59-0.66)	0.57 (0.53-0.61)	1053	68

Table-A I-3 Accuracy (< 5 mmHg) comparison in hypo-, normo- and hypercapnia

Time gap with calibration ABG	< 35 mmHg	[35-45] mmHg	> 45 mmHg	Global Accuracy	
2h -> Fold nb	1	24%	77%	51%	59%
	2	23%	83%	52%	63%
	3	36%	74%	61%	62%
	4	32%	73%	48%	57%
	5	40%	81%	44%	61%
Average	31%	78%	51%	60%	
6h -> Fold nb	1	33%	81%	46%	64%
	2	39%	77%	52%	61%
	3	37%	79%	54%	63%
	4	40%	77%	52%	63%
	5	43%	81%	54%	66%
Average	38%	79%	51%	64%	
12h -> Fold nb	1	20%	77%	54%	63%
	2	26%	78%	56%	63%
	3	25%	84%	51%	65%
	4	33%	72%	54%	60%
	5	14%	73%	55%	60%
Average	24%	77%	54%	62%	
24h -> Fold nb	1	15%	76%	66%	67%
	2	25%	79%	63%	68%
	3	30%	79%	60%	66%
	4	20%	85%	59%	69%
	5	33%	76%	62%	66%
Average	25%	79%	62%	67%	

Table-A I-4 Bland Altman Characteristics

Bias, upper and lower 95% limits of agreement and their 95% confidence intervals

Time gap with calibration ABG	Bias	Upper limit	Lower limit	
2h → Fold nb	1	-0.73 ±0.81	15.15 ±1.39	-16.61 ±1.39
	2	1.13 ±0.76	16.04 ±1.30	-13.78 ±1.30
	3	-0.39 ±0.70	13.56 ±1.20	-14.35 ±1.20
	4	-0.31 ±0.85	16.51 ±1.46	-17.13 ±1.46
	5	1.25 ±0.90	19.72 ±1.55	-17.22 ±1.55
	Average	0.19 ±0.80	16.20 ±1.38	-15.82 ±1.38
6h → Fold nb	1	0.05 ±0.45	14.23 ±0.76	-14.13 ±0.76
	2	0.36 ±0.48	16.09 ±0.82	-15.37 ±0.82
	3	0.64 ±0.42	14.04 ±0.72	-12.76 ±0.72
	4	-0.04 ±0.38	11.99 ±0.65	-12.07 ±0.65
	5	0.38 ±0.40	13.15 ±0.68	-12.40 ±0.68
	Average	0.28 ±0.43	13.90 ±0.73	-13.35 ±0.73
12h → Fold nb	1	0.27 ±0.41	15.15 ±0.70	-14.60 ±0.70
	2	0.11 ±0.37	13.24 ±0.63	-13.02 ±0.63
	3	0.55 ±0.33	12.57 ±0.57	-11.46 ±0.57
	4	-0.16 ±0.41	14.76 ±0.70	-15.08 ±0.70
	5	0.95 ±0.41	15.65 ±0.69	-13.75 ±0.69
	Average	0.34 ±0.38	14.27 ±0.66	-13.58 ±0.66
24h → Fold nb	1	0.00 ±0.39	13.07 ±0.67	-13.07 ±0.67
	2	0.27 ±0.36	11.87 ±0.62	-11.34 ±0.62
	3	-0.18 ±0.40	12.55 ±0.68	-12.90 ±0.68
	4	0.13 ±0.35	11.63 ±0.60	-11.37 ±0.60
	5	-0.18 ±0.40	12.87 ±0.68	-13.22 ±0.68
	Average	0.01 ±0.38	12.40 ±0.65	-12.38 ±0.65

ANNEXE II

RÉSULTATS DES 5 FOLDS POUR LES MODÈLES MLP + PETCO₂

Table-A II-1 Accuracy (≤ 5 mmHg) robustness comparison when significant changes occur between calibration ABG and predicted PaCO₂

Time gap with calibration ABG	Accuracy if PaCO₂ was unchanged	Accuracy if PaCO₂ changed
2h → Fold nb	1	89%
	2	83%
	3	82%
	4	87%
	5	87%
Average	86%	61%
6h → Fold nb	1	90%
	2	95%
	3	100%
	4	91%
	5	81%
Average	91%	47%
12h → Fold nb	1	93%
	2	83%
	3	85%
	4	87%
	5	90%
Average	87%	63%
24h → Fold nb	1	74%
	2	90%
	3	91%
	4	93%
	5	100%
Average	90%	67%

Table-A II-2 Prediction validation

Time gap with calibration ABG	MedAE (IQR)	R2 (95% CI)	ICC (95% CI)	Number of ABGs used for testing	Number of patients
2h → Fold nb	1 3.75 (2.6-4.9)	0.91 (0.79-0.96)	0.87 (0.72-0.94)	22	11
	2 3.6 (2.1-4.9)	0.69 (0.41-0.85)	0.67 (0.39-0.84)	25	11
	3 2.2 (1.1-4.0)	0.93 (0.84-0.97)	0.93 (0.84-0.97)	26	12
	4 3.4 (1.7-5.5)	0.61 (0.27-0.81)	0.6 (0.28-0.8)	24	12
	5 3.9 (2.3-5.6)	0.71 (0.44-0.86)	0.69 (0.43-0.85)	26	11
Average	3.37 (2.0-5.0)	0.77 (0.55-0.89)	0.752 (0.53-0.88)	25	11
6h → Fold nb	1 2.8 (1.7-4.9)	0.75 (0.61-0.85)	0.73 (0.58-0.83)	56	22
	2 3 (1.9-4.4)	0.75 (0.61-0.84)	0.7 (0.55-0.81)	58	21
	3 2.6 (1.4-4.4)	0.71 (0.56-0.82)	0.69 (0.53-0.80)	59	21
	4 3.01 (1.6-4.7)	0.82 (0.71-0.89)	0.78 (0.66-0.86)	60	21
	5 3.2 (2.0-7.5)	0.63 (0.46-0.76)	0.51 (0.30-0.67)	64	22
Average	2.922 (1.7-5.2)	0.732 (0.59-0.83)	0.682 (0.52-0.79)	59	21
12h → Fold nb	1 2.7 (1.2-4.4)	0.85 (0.77-0.91)	0.83 (0.73-0.89)	61	23
	2 3.5 (1.8-6.2)	0.67 (0.51-0.79)	0.58 (0.38-0.72)	62	25
	3 2.8 (1.0-4.8)	0.84 (0.75-0.9)	0.82 (0.71-0.89)	60	25
	4 3.1 (1.5-5.0)	0.87 (0.79-0.92)	0.77 (0.64-0.85)	61	24
	5 3.1 (1.8-4.6)	0.78 (0.66-0.86)	0.75 (0.62-0.84)	61	21
Average	3.04 (1.5-5.0)	0.802 (0.70-0.88)	0.75 (0.62-0.84)	61	24

Table-A II-2 Prediction validation (continued)

Time gap with calibration ABG	MedAE (IQR)	R2 (95% CI)	ICC (95% CI)	Number of ABGs used for testing	Number of patients
<i>24h -> Fold nb</i> 1	3.4 (2.3-6.4)	0.81 (0.71-0.88)	0.76 (0.63-0.84)	65	21
2	3 (1.8-4.7)	0.67 (0.51-0.79)	0.66 (0.5-0.78)	61	20
3	3.2 (1.5-4.8)	0.66 (0.49-0.78)	0.62 (0.44-0.75)	63	22
4	2.6 (1.3-4.8)	0.64 (0.46-0.77)	0.6 (0.42-0.74)	63	21
5	2.7 (1.4-4.2)	0.66 (0.5-0.78)	0.65 (0.49-0.77)	66	20
<i>Average</i>	2.98 (1.6-5.0)	0.688 (0.54-0.8)	0.658 (0.50-0.78)	64	21

Table-A II-3 Accuracy (< 5 mmHg) comparison in hypo-, normo- and hypercapnia

Time gap with calibration ABG	< 35 mmHg	[35-45] mmHg	> 45 mmHg	Global Accuracy	
2h → Fold nb	1	91%	78%	50%	82%
	2	86%	60%	88%	76%
	3	89%	91%	67%	85%
	4	100%	78%	20%	75%
	5	50%	85%	67%	73%
Average	83%	78%	58%	78%	
6h → Fold nb	1	43%	88%	65%	75%
	2	89%	92%	50%	84%
	3	89%	90%	68%	83%
	4	75%	86%	73%	82%
	5	71%	88%	40%	67%
Average	73%	89%	59%	78%	
12h → Fold nb	1	44%	91%	94%	85%
	2	57%	95%	61%	73%
	3	50%	92%	72%	78%
	4	100%	96%	62%	77%
	5	50%	88%	77%	80%
Average	60%	93%	73%	79%	
24h → Fold nb	1	33%	95%	65%	72%
	2	71%	91%	71%	82%
	3	63%	91%	70%	79%
	4	60%	82%	74%	76%
	5	75%	94%	86%	88%
Average	60%	91%	73%	80%	

Table-A II-4 Bland Altman Characteristics
 Bias, upper and lower 95% limits of agreement and their 95% confidence intervals

Time gap with calibration ABG	Bias	Upper limit	Lower limit
2h → Fold nb	1	-2.69 ±1.66	5.08 ±3.05
	2	-2.60 ±2.97	12.26 ±5.42
	3	0.09 ±1.36	7.00 ±2.47
	4	-0.40 ±3.68	17.65 ±6.74
	5	-1.66 ±2.40	10.56 ±4.36
	Average	-1.45 ±2.41	10.51 ±4.41
6h → Fold nb	1	0.25 ±1.65	12.61 ±2.90
	2	0.70 ±1.27	10.36 ±2.23
	3	0.63 ±1.63	13.17 ±2.86
	4	-0.61 ±1.25	9.05 ±2.19
	5	3.28 ±1.95	18.92 ±3.42
	Average	0.85 ±1.55	12.82 ±2.72
12h → Fold nb	1	-0.65 ±1.07	7.74 ±1.88
	2	2.27 ±1.84	16.75 ±3.22
	3	0.96 ±1.15	9.90 ±2.03
	4	3.05 ±1.43	14.24 ±2.51
	5	1.01 ±1.24	10.71 ±2.18
	Average	1.33 ±1.35	11.87 ±2.36
24h → Fold nb	1	1.94 ±1.42	13.43 ±2.49
	2	0.44 ±1.34	10.88 ±2.34
	3	-0.89 ±1.20	8.66 ±2.11
	4	-0.66 ±1.45	10.88 ±2.55
	5	-0.72 ±1.17	8.82 ±2.05
	Average	0.02 ±1.32	10.53 ±2.31

ANNEXE III

LES MEILLEURES COMBINAISONS SELON LES DIFFÉRENTS SOUS GROUPES DANS LA CATÉGORIE 2H

Table-A III-1 Les meilleures combinaisons

Unchanged PaCO ₂				
Combinaison	< 35 mmHg	[35-45] mmHg	> 45 mmHg	Global Accuracy
1	93%	95%	88%	92%
2	83%	99%	84%	91%
3	72%	97%	93%	91%
1	93%	95%	88%	92%
Changed PaCO ₂				
Combinaison	< 35 mmHg	[35-45] mmHg	> 45 mmHg	Global Accuracy
4	26%	37%	29%	30%
5	0%	88%	25%	37%
6	0%	45%	55%	43%
6	0%	45%	55%	43%
All PaCO ₂				
Combinaison	< 35 mmHg	[35-45] mmHg	> 45 mmHg	Global Accuracy
7	50%	64%	39%	50%
8	2%	79%	28%	42%
9	25%	57%	57%	51%
10	34%	71%	54%	56%

Table-A III-2 Parameters distribution on the different combinations

Parameters	Combination									
	1	2	3	4	5	6	7	8	9	10
PaCO ₂	x	x	x	x			x		x	x
RR	x			x			x		x	x
FiO ₂				x	x			x	x	x
Vte						x			x	
PIP	x	x	x	x	x	x		x		x
rap			x	x	x				x	
MawP	x	x		x		x	x		x	x
Vmin				x						x
Time Gap				x		x			x	x

RÉFÉRENCES

- Adel Zaqoot, H., Baloch, A., Khaliq Ansari, A. & Ali Unar, M. (2010). Application of artificial neural networks for predicting pH in seawater along Gaza beach. *Applied Artificial Intelligence*, 24(7), 667–679.
- Baraka, A., El-Khatib, M., Muallem, E., Jamal, S., Haroun-Bizri, S. & Aouad, M. (2005). Oxygenator exhaust capnography for prediction of arterial carbon dioxide tension during hypothermic cardiopulmonary bypass. *The journal of extra-corporeal technology*, 37(2), 192.
- Baudin, F., Bourgoin, P., Brossier, D., Essouri, S., Emeriaud, G., Wysocki, M. & Jouvét, P. (2016). Noninvasive Estimation of Arterial CO₂ From End-Tidal CO₂ in Mechanically Ventilated Children : The GRAeDIENT Pilot Study. *Pediatric Critical Care Medicine*, 17(12), 1117–1123.
- Bhaduri, A., Bhaduri, A., Bhaduri, A. & Mohapatra, P. (2009). Blood pressure modeling using statistical and computational intelligence approaches. *2009 IEEE International Advance Computing Conference*, pp. 1026–1030.
- Bland, J. M. & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476), 307–310.
- Brossier, D., El Taani, R., Sauthier, M., Roumeliotis, N., Emeriaud, G. & Jouvét, P. (2018). Creating a high-frequency electronic database in the PICU : the perpetual patient. *Pediatric critical care medicine*, 19(4), e189–e198.
- Brothers, R. M., Ganio, M. S., Hubing, K. A., Hastings, J. L. & Crandall, C. G. (2011). End-tidal carbon dioxide tension reflects arterial carbon dioxide tension in the heat-stressed human with and without simulated hemorrhage. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 300(4), R978–R983.
- Chan, K., Chan, M. & Gin, T. (2003). Mainstream vs. sidestream capnometry for prediction of arterial carbon dioxide tension during supine craniotomy. *Anaesthesia*, 58(2), 149–155.
- Choubin, B., Khalighi-Sigaroodi, S., Malekian, A. & Kişi, Ö. (2016). Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals. *Hydrological Sciences Journal*, 61(6), 1001–1009.
- Eberhard, P. (2007). The design, use, and results of transcutaneous carbon dioxide analysis : current and future directions. *Anesthesia & Analgesia*, 105(6), S48–S52.
- Fearon, D. M. & Steele, D. W. (2002). End-tidal carbon dioxide predicts the presence and severity of acidosis in children with diabetes. *Academic Emergency Medicine*, 9(12),

1373–1378.

- Gardner, M. W. & Dorling, S. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*, 32(14-15), 2627–2636.
- Gelsomino et al. (2011). Assessment of a continuous blood gas monitoring system in animals during circulatory stress. *BMC anesthesiology*, 11(1), 1.
- Glorot, X. & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 249–256.
- Haykin, S. (1998). *Neural Networks : A Comprehensive Foundation* (éd. 2nd). Upper Saddle River, NJ, USA : Prentice Hall PTR.
- Jaffe, M. B. (2017). Using the features of the time and volumetric capnogram for classification and prediction. *Journal of clinical monitoring and computing*, 31(1), 19–41.
- Jakkula, P., Reinikainen, M., Hästbacka, J., Loisa, P., Tiainen, M., Pettilä, V., Toppila, J., Lähde, M., Bäcklund, M., Okkonen, M. et al. (2018). Targeting two different levels of both arterial carbon dioxide and arterial oxygen after cardiac arrest and resuscitation : a randomised pilot trial. *Intensive care medicine*, 44(12), 2112–2121.
- Khemani, R. G., Celikkaya, E. B., Shelton, C. R., Kale, D., Ross, P. A., Wetzel, R. C. & Newth, C. J. (2014). Algorithms to estimate PaCO₂ and pH using noninvasive parameters for children with hypoxemic respiratory failure. *Respiratory care*, 59(8), 1248–1257.
- Khemani, R. G., Smith, L. S., Zimmerman, J. J., Erickson, S., Group, P. A. L. I. C. C. et al. (2015). Pediatric acute respiratory distress syndrome : definition, incidence, and epidemiology : proceedings from the Pediatric Acute Lung Injury Consensus Conference. *Pediatric Critical Care Medicine*, 16(5_suppl), S23–S40.
- Kingma, D. & Ba, J. (2014). Adam : A Method for Stochastic Optimization. *International Conference on Learning Representations*.
- Kirk, V. G., Batuyong, E. D. & Bohn, S. G. (2006). Transcutaneous carbon dioxide monitoring and capnography during pediatric polysomnography. *Sleep*, 29(12), 1601–1608.
- Koo, T. K. & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, 15(2), 155–163.
- Lermuzeaux, M., Meric, H., Sauneuf, B., Girard, S., Normand, H., Lofaso, F. & Terzi, N. (2016). Superiority of transcutaneous CO₂ over end-tidal CO₂ measurement for monitoring respiratory failure in nonintubated patients : A pilot study. *Journal of critical care*, 31(1),

150–156.

- Martinez-Coll, A. & Nguyen, H. (2001). Comparison of near infrared spectroscopy (NIRS) signal quantitation by multilinear regression and neural networks. *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2, 1625–1628.
- McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115–133.
- McSwain, S. D., Hamel, D. S., Smith, P. B., Gentile, M. A., Srinivasan, S., Meliones, J. N. & Cheifetz, I. M. (2010). End-tidal and arterial carbon dioxide measurements correlate across all levels of physiologic dead space. *Respiratory care*, 55(3), 288–293.
- Medicare, M. (1992). CLIA programs; regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA)-HCFA. Final rule with comment period. *Fed Regist*, 57(40), 7002–7186.
- Naeini, M. P., Taremian, H. & Hashemi, H. B. (2010). Stock market value prediction using neural networks. *2010 international conference on computer information systems and industrial management applications (CISIM)*, pp. 132–136.
- Organization, W. H. et al. (2010). *WHO guidelines on drawing blood : best practices in phlebotomy*. World Health Organization.
- Panigrahi, N., Ahmed, M., Mishra, M. & V, B. (2018). Predictive capability of End Tidal carbon dioxide and its correlation with arterial carbon dioxide in mechanically ventilated children. 5, 15. doi : 10.21304/2018.0504.00404.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn : Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Potger, K. C., McMillan, D., Southwell, J., Dando, H. & O Shaughnessy, K. (2003). Membrane oxygenator exhaust capnography for continuously estimating arterial carbon dioxide tension during cardiopulmonary bypass. *Journal of ExtraCorporeal Technology*, 35(3), 218–223.
- Razi, E., Moosavi, G. A., Omidi, K., Saebi, A. K. & Razi, A. (2012). Correlation of end-tidal carbon dioxide with arterial carbon dioxide in mechanically ventilated patients. *Archives of trauma research*, 1(2), 58.
- Rentola, R., Hästbacka, J., Heinonen, E., Rosenberg, P. H., Häggblom, T. & Skrifvars, M. B. (2018). Estimation of Arterial Carbon Dioxide Based on End-Tidal Gas Pressure and

Oxygen Saturation. *Journal of clinical medicine*, 7(9), 290.

Rumelhart, D. E., Hinton, G. E., Williams, R. J. et al. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, 5(3), 1.

Sasse, S. A., Chen, P. A. & Mahutte, C. K. (1994). Variability of arterial blood gas values over time in stable medical ICU patients. *Chest*, 106(1), 187–193.

Shan, L. & Jie, X. (2011). Artificial neural networks prediction for blood concentration and dosage of cyclosporine A. *2011 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1875–1880.

Shenouda, E. A. A. (2006). A quantitative comparison of different MLP activation functions in classification. *International Symposium on Neural Networks*, pp. 849–857.

Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations : uses in assessing rater reliability. *Psychological bulletin*, 86(2), 420.

Sood, P., Paul, G. & Puri, S. (2010). Interpretation of arterial blood gas. *Indian journal of critical care medicine : peer-reviewed, official publication of Indian Society of Critical Care Medicine*, 14(2), 57.

Stammers, A. H., Mejak, B. L., Rauch, E. D., Vang, S. N. & Viessman, T. W. (2000). Factors affecting perfusionists' decisions on equipment utilization : results of a United States survey. *JOURNAL OF EXTRACORPOREAL TECHNOLOGY*, 32(1), 4–10.

Vaahersalo, J., Bendel, S., Reinikainen, M., Kuroola, J., Tiainen, M., Raj, R., Pettilä, V., Varpula, T., Skrifvars, M. B., Group, F. S. et al. (2014). Arterial blood gas tensions after resuscitation from out-of-hospital cardiac arrest : associations with long-term neurologic outcome. *Critical care medicine*, 42(6), 1463–1470.

Wexler, H. R. & Lok, P. (1981). A simple formula for adjusting arterial carbon dioxide tension. *Canadian Anaesthetists' Society Journal*, 28(4), 370–372. doi : 10.1007/BF03007805.

Zernikow, B., Holtmannspoetter, K., Michel, E., Theilhaber, M., Pielemeier, W. & Hennecke, K. (1998). Artificial neural network for predicting intracranial haemorrhage in preterm neonates. *Acta Paediatrica*, 87(9), 969–975.