Uncertainty analysis, sensitivity analysis, and machine learning in cardiovascular biomechanics

Robert Rapadamnaba

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR
DE L'UNIVERSITÉ DE MONTPELLIER

En Mathématiques et modélisation

École doctorale I2S - Information, Structures, Systèmes

Unité de recherche UMR 5149 - IMAG - Institut Montpelliérain Alexander Grothendieck

Uncertainty analysis, sensitivity analysis, and machine learning in cardiovascular biomechanics

Présentée par Robert RAPADAMNABA

Le 25 septembre 2020

Sous la direction de Bijan MOHAMMADI et Franck NICOUD

Devant le jury composé de

Bruno KOOBUS	Professeur	Université de Montpellier	Président du jury
José-Maria FULLANA	Professeur	Université Pierre et Marie Curie	Examinateur
Bijan MOHAMMADI	Professeur	Université de Montpellier	Directeur
Franck NICOUD	Professeur	Université de Montpellier	Co-Directeur
Christophe PRUD'HOMME	Professeur	Université de Strasbourg	Rapporteur
Irène VIGNON-CLEMENTEL	Directrice de recherche	INRIA de Paris	Rapporteure

UNIVERSITÉ DE MONTPELLIER

Uncertainty analysis, sensitivity analysis, and machine learning in cardiovascular biomechanics

Abstract:

This thesis follows on from a recent study conducted by a few researchers from University of Montpellier, with the aim of proposing to the scientific community an inversion procedure capable of noninvasively estimating patient-specific blood pressure in cerebral arteries. Its first objective is, on the one hand, to examine the accuracy and robustness of the inversion procedure proposed by these researchers with respect to various sources of uncertainty related to the models used, formulated assumptions and patient-specific clinical data, and on the other hand, to set a stopping criterion for the ensemble Kalman filter based algorithm used in their inversion procedure. For this purpose, uncertainty analysis and several sensitivity analyses are carried out. The second objective is to illustrate how machine learning, mainly focusing on convolutional neural networks, can be a very good alternative to the time-consuming and costly inversion procedure implemented by these researchers for cerebral blood pressure estimation.

An approach taking into account the uncertainties related to the patient-specific medical images processing and the blood flow model assumptions, such as assumptions about boundary conditions, physical and physiological parameters, is first presented to quantify uncertainties in the inversion procedure outcomes. Uncertainties related to medical images segmentation are modelled using a Gaussian distribution and uncertainties related to modeling assumptions choice are analyzed by considering several possible hypothesis choice scenarii. From this approach, it emerges that the uncertainties on the procedure results are of the same order of magnitude as those related to segmentation errors. Furthermore, this analysis shows that the procedure outcomes are very sensitive to the assumptions made about the model boundary conditions. In particular, the choice of the symmetrical Windkessel boundary conditions for the model proves to be the most relevant for the case of the patient under study.

Next, an approach for ranking the parameters estimated during the inversion procedure in order of importance and setting a stopping criterion for the algorithm used in the inversion procedure is presented. The results of this strategy show, on the one hand, that most of the model proximal resistances are the most important parameters for blood flow estimation in the internal carotid arteries and, on the other hand, that the inversion algorithm can be stopped as soon as a certain reasonable convergence threshold for the most influential parameter is reached.

Finally, a new numerical platform, based on machine learning and allowing to estimate the patient-specific blood pressure in the cerebral arteries much faster than with the inversion procedure but with the same accuracy, is presented. The application of this platform to the patient-specific data used in the inversion procedure provides noninvasive and real-time estimate of patient-specific cerebral pressure consistent with the inversion procedure estimation.

Keywords: inverse problem, uncertainty analysis, sensitivity analysis, machine learning, convolutional neural networks, ensemble Kalman filter, 0D compartment model, hemodynamics, circle of Willis, cerebral aneurysms, patient-specific simulation, blood pressure estimation.

Analyse d'incertitude, Analyse de Sensibilité et Apprentissage machine en Biomécanique Cardiovasculaire

Résumé

Cette thèse fait suite à une étude récente, menée par quelques chercheurs de l'Université de Montpellier, dans le but de proposer à la communauté scientifique une procédure d'inversion capable d'estimer de manière non invasive la pression dans les artères cérébrales d'un patient. Son premier objectif est, d'une part, d'examiner la précision et la robustesse de la procédure d'inversion proposée par ces chercheurs, en lien avec diverses sources d'incertitude liées aux modèles utilisés, aux hypothèses formulées et aux données cliniques du patient, et d'autre part, de fixer un critère d'arrêt pour l'algorithme basé sur le filtre de Kalman d'ensemble utilisé dans leur procédure d'inversion. À cet effet, une analyse d'incertitude et plusieurs analyses de sensibilité sont effectuées. Le second objectif est d'illustrer comment l'apprentissage machine, orienté réseaux de neurones convolutifs, peut être une très bonne alternative à la longue et coûteuse procédure mise en place par ces chercheurs pour l'estimation de la pression.

Une approche prenant en compte les incertitudes liées au traitement des images médicales du patient et aux hypothèses formulées sur les modèles utilisés, telles que les hypothèses liées aux conditions limites, aux paramètres physiques et physiologiques, est d'abord présentée pour quantifier les incertitudes sur les résultats de la procédure. Les incertitudes liées à la segmentation des images sont modélisées à l'aide d'une distribution gaussienne et celles liées au choix des hypothèses de modélisation sont analysées en testant plusieurs scénarios de choix d'hypothèses possibles. De cette démarche, il ressort que les incertitudes sur les résultats de la procédure sont du même ordre de grandeur que celles liées aux erreurs de segmentation. Par ailleurs, cette analyse montre que les résultats de la procédure sont très sensibles aux hypothèses faites sur les conditions aux limites du modèle du flux sanguin. En particulier, le choix des conditions limites symétriques de Windkessel pour le modèle s'avère être le plus approprié pour le cas du patient étudié.

Ensuite, une démarche permettant de classer les paramètres estimés à l'aide de la procédure par ordre d'importance et de fixer un critère d'arrêt pour l'algorithme utilisé dans cette procédure est proposée. Les résultats de cette stratégie montrent, d'une part, que la plupart des résistances proximales sont les paramètres les plus importants du modèle pour l'estimation du débit sanguin dans les carotides internes et, d'autre part, que l'algorithme d'inversion peut être arrêté dès qu'un certain seuil de convergence raisonnable de ces paramètres les plus influents est atteint.

Enfin, une nouvelle plateforme numérique basée sur l'apprentissage machine permettant d'estimer la pression artérielle spécifique au patient dans les artères cérébrales beaucoup plus rapidement qu'avec la procédure d'inversion mais avec la même précision, est présentée. L'application de cette plateforme aux données du patient utilisées dans la procédure d'inversion permet une estimation non invasive et en temps réel de la pression dans les artères cérébrales du patient cohérente avec l'estimation de la procédure d'inversion.

Mots-clés: problème inverse, analyse d'incertitude, analyse de sensibilité, apprentissage machine, réseaux de neurones convolutifs, filtre de Kalman d'ensemble, modèle réduit 0D, hémodynamique, polygone de Willis, anévrismes cérébraux, simulation spécifique au patient, estimation de la pression sanguine.

"Hiding within those mounds of data is knowledge that could change the life of a patient, or change the world."

Atul Butte

"You can have data without information, but you cannot have information without data."

Daniel Keys Moran

"Torture the data, and it will confess to anything"

Ronald Coase

To my parents, my brothers and my sister,
and above all, to my lovely little family: Joséphine, Graciella, and Dylan.

Remerciements

Il me sera très difficile en ces quelques lignes de remercier individuellement toutes les nombreuses personnes par qui cette thèse a vu le jour. Cependant, je m'en voudrais de ne pas nommer quelques unes d'entre elles grâce à qui j'ai pu mener cette thèse à son terme.

Je voudrais, en premier lieu, exprimer toute ma reconnaissance à mes deux encadrants de thèse: Pr. **Bijan MOHAMMADI** et Pr. **Franck NICOUD**. Merci de m'avoir fait confiance en me proposant un sujet de recherche aussi original que passionnant; merci de votre constante et précieuse aide durant ces trois années denses et intenses. J'ai énormément appris à vos côtés et vous ne vous êtes jamais échappés face à mes questions et doutes de jeune chercheur. Au contraire, vous avez toujours été là pour me soutenir, me motiver, m'inspirer et me conseiller au cours de l'élaboration de cette thèse. J'ai vraiment été ravi d'avoir pu bénéficier de vos expériences respectives et de votre appui scientifique au cours de ces années. Merci aussi pour la bonne ambiance qui a toujours régné lors de nos réunions à deux, à trois, ou avec toute l'équipe `Yales2bio`. J'en suis tellement comblé que si l'on me demandait aujourd'hui, après ces trois belles années passées à votre école, de refaire une autre thèse avec vous deux comme encadrants, je répondrais sans aucune hésitation par l'affirmative.

Je remercie également Dr. **Irene VIGNON-CLEMENTEL** et Pr. **Christophe PRUD'HOMME** qui m'ont fait l'honneur d'être rapporteurs de ma thèse. Vous avez sacrifié vos vacances d'été pour la lecture de ce manuscrit et vous avez pris le temps de m'écouter et de discuter avec moi. Vos remarques et questions pertinentes m'ont permis d'améliorer la qualité de ce manuscrit et d'envisager pour le futur de nouvelles perspectives de recherche très prometteuses. Merci à vous! Un grand merci aussi à Dr. **Emmanuelle LE BARS** pour m'avoir permis d'obtenir l'imagerie médicale nécessaire à la réalisation de cette thèse, pour son apport scientifique, son expertise médicale et pour avoir accepté de participer à mon jury de thèse. A travers vous, j'exprime ma profonde gratitude à tout le Département de Neuroradiologie du CHRU Gui de Chauliac de Montpellier pour la fructueuse collaboration durant ces années de thèse. J'ai également apprécié la disponibilité, le temps et l'intérêt qu'ont accordés les Professeurs

José-Maria FULLANA et **Bruno KOOBUS** à mon travail en participant à mon jury de thèse. Merci à tous pour le travail colossal abattu!

Je tiens aussi à exprimer ma gratitude envers les personnes et institutions sans lesquelles cette thèse n'aurait jamais pu être possible: **Mgr Thomas KABORÉ**, évêque émérite de Kaya, les **diocèses de Kaya et de Nîmes**, le **père Jacques CHANUT** et à travers lui, tout **l'archidiocèse de Montpellier**, et par-dessus tout **l'État français** qui non seulement m'a donné la chance de pouvoir accéder librement à la connaissance en suivant des études supérieures dans de bonnes conditions, mais aussi m'a énormément soutenu en finançant ces années de thèse.

Je ne pourrais pas non plus passer sous silence tous ceux sans qui cette thèse ne serait pas ce qu'elle est aujourd'hui, aussi bien par les discussions que j'ai eu la chance d'avoir avec eux que par leurs suggestions, conseils ou contributions. Je pense ici en particulier à Dr. **Simon MENDEZ**, Dr. **Thierry MIGNON** et à tous mes collègues de l'équipe `Yales2bio`: Dr. **Rajnesh LAL**, Dr. **Stéphanie LINDSEY**, Dr. **Thomas PUISEUX**, Dr. **Pierre TARACONAT**, Dr. **Dorian MIDOU**, Dr. **Rodrigo MÉNDEZ ROJANO**, **Pascal MATTEOLI**, **Alain BEROD**, **Samuel ORRU**, **Agathe BIGNON** et **Morgane GARREAU**. Un merci spécial à **Pascal MATTEOLI** qui s'est toujours rendu disponible pour m'apporter l'aide nécessaire à l'avancée de ma thèse. "*N'hésite pas hein!!! si tu as besoin...*" m'as-tu toujours répété, Pascal. Alors, vrai gars, mon poto, reçois ici l'expression de ma profonde gratitude!

Il m'est également impossible de ne pas nommer individuellement chacun de mes co-bureaux: Dr. **Néstor Fernández VARGAS**, **Abdoulrahim IBRAHIM YOUSSOUF**, **Gwenaël PELTIER** et **André HARNIST**. Les gars! Ce fut un honneur pour moi de partager le **bureau 115** avec vous à l'Institut Montpelliérain Alexander Grothendieck. Merci pour les échanges scientifiques ou non que nous avons pu avoir et pour l'ambiance générale qui a régné durant ces années de thèse. *Big up à chacun de vous!*

Un grand merci aux **personnels administratifs et techniques du laboratoire** et à mes **collègues étudiantes, étudiants, doctorantes et doctorants** avec qui j'ai partagé mes années d'études et notamment ces années de thèse. J'ai pu côtoyer la plupart d'entre vous, que ce soit pour des discussions et événements scientifiques ponctuels autour et au-delà de la thèse, ou simplement pour des moments conviviaux et j'en garde un très bon souvenir. Ce fut un réel plaisir d'avoir fait un bout de chemin avec vous.

J'ai ici aussi une pensée spéciale pour mes **parents**, mes **frères** et ma **petite sœur**. Je n'aurais jamais réussi ce long périple sans vos bénédictions et vos prières. Alors merci pour tout!

Je salue surtout avec une immense et profonde gratitude l'amour, les

encouragements, le soutien continuel et la foi inébranlable en moi que ma moitié **Joséphine Marie Joseph Suutongnooma** m'a donnés durant ce long voyage de thèse. Mon cœur, ta patience et ton sacrifice légendaires resteront éternellement gravés en moi. Tu as consenti à d'énormes sacrifices et tu as eu la force, le courage et la patience de supporter et de porter toute seule sur tes épaules, trop de peines et de charges. Tu t'es occupée de tout ce qui était nécessaire sans te plaindre, juste pour que je puisse me concentrer sur mes études et sur la réalisation de ma thèse. Oui! Tu as patiemment enduré de nombreuses et longues années de séparation et de solitude pendant que je me consacrais à mes études et à ma thèse. Je te remercie de tout mon cœur et de toute mon âme. Ton amour total et inconditionnel me portera toujours à bout de bras. Ces remerciements ne seraient pas complets si je ne mentionnais pas mes adorables jumeaux de 6 ans, **Marie-Josée Graciella et José Dylan Auguste**, qui sont mes joyaux depuis leur naissance. Ce fut très douloureux de ne pas vous voir grandir à mes côtés mais vos sourires m'ont encouragé à surmonter efficacement toutes les difficultés que j'ai rencontrées. Je vous exprime mon amour et ma reconnaissance les plus profonds. Je vous dois à tous les trois beaucoup beaucoup d'heures de loisirs et de plaisir.

Mes derniers remerciements vont aux personnes de l'ombre qui n'aimeraient pas être nommées explicitement. Toutes ces petites mains qui ont tout fait pour m'aider, me soutenir et m'encourager dans tout ce que j'ai entrepris durant ces années. Veuillez trouver dans ces quelques lignes l'expression de ma sincère gratitude.

À chacun:
Merci beaucoup, Thank you so much, Y bark barka.

R. RAPADAMNABA

Préface

Les maladies cardiovasculaires sont devenues un problème de santé publique de nos jours. En effet, selon l'Organisation mondiale de la santé (OMS), les maladies cardiovasculaires sont la principale cause de mortalité dans le monde, avec 17,9 millions de morts chaque année, soit 31 % de la mortalité mondiale. Parmi ces maladies, les anévrismes cérébraux qui touchent 3,6 à 6% de la population mondiale, nécessitent une attention particulière car ils comportent un risque élevé de mortalité en cas de rupture. En effet, une fois qu'un anévrisme cérébral saigne ou se rompt, le sang se répand dans l'espace sous-arachnoïdien (l'espace entre le crâne et le cerveau), provoquant une hémorragie sous-arachnoïdienne [1, 2], qui peut rapidement s'avérer très fatale dans la plupart du temps et des cas, avec 40 % de risque de décès [3, 4, 5].

L'un des principaux facteurs identifiés comme étant associé à la formation et au risque de rupture des anévrismes cérébraux est la fluctuation de la pression sanguine dans les artères cérébrales [6, 7, 8, 9, 10, 11]. Par conséquent, une technique permettant d'estimer ces variations pressionnelles dans les artères cérébrales, de manière non invasive, serait bénéfique et d'un grand intérêt pour la prise de décision clinique et le traitement des anévrismes cérébraux. C'est dans ce contexte que quelques chercheurs de l'Université de Montpellier ont récemment mis en œuvre une plateforme d'inversion basée sur le filtre de Kalman d'ensemble couplé à un modèle d'ordre réduit 0D, dans le but de proposer à la communauté scientifique une technique d'estimation non invasive de la pression, dans les artères cérébrales d'un patient [12]. Cette procédure d'inversion, qui implique des acquisitions d'imagerie médicale du patient, utilise un réseau artériel complexe constitué de 33 artères incluant le polygone de Willis (i.e., les artères situées à la base du cerveau) du patient. Toutefois, la précision et la robustesse de cette plateforme d'inversion par rapport à certaines hypothèses de modélisation doivent encore être examinées et nécessitent donc plus d'investigations. En outre, cette procédure d'inversion qui nécessite la résolution d'un problème inverse, est trop longue (aucun critère d'arrêt n'y a été fixé), coûteuse,

difficile à intégrer dans un dispositif d'acquisition d'imagerie médicale, et ses résultats restent encore à valider. Ce sont donc autant de pistes et de centres d'intérêts qui méritent d'être creusés et dont l'exploration permet d'apporter des éléments de réponse aux questions suivantes:

- Quel est l'impact des différentes sources d'incertitude liées au modèles utilisés, aux hypothèses formulées ainsi qu'aux données cliniques du patient sur les résultats de leur procédure d'inversion?

- Les incertitudes sur les résultats de la procédure peuvent-elles être quantifiées et comment? Si oui, quel est alors le degré de fiabilité des résultats de la plateforme?

- Est-il possible de fixer un critère d'arrêt pour l'algorithme utilisé dans la procédure d'inversion pour éviter une sur-résolution inutile et réduire ainsi le temps de calcul?

- Comment accélérer cette procédure d'estimation de la pression, en particulier dans les situations cliniques où un diagnostic rapide est toujours souhaité?

Continuer le travail entrepris par Lal et al. [12] en répondant efficacement à ces différentes questions est essentiellement le défi, l'ambition et l'enjeu principal de cette thèse. En particulier, son premier objectif est, d'une part, d'examiner la précision et la robustesse de la procédure d'inversion proposée par ces chercheurs, en lien avec diverses sources d'incertitude liées aux modèles, aux hypothèses formulées et aux données cliniques utilisées, et d'autre part, de fixer un critère d'arrêt pour l'algorithme basé sur le filtre de Kalman d'ensemble utilisé dans leur procédure d'inversion. À cet effet, une analyse d'incertitude et plusieurs analyses de sensibilité sont effectuées. Le second objectif majeur est d'illustrer comment l'apprentissage machine, orienté réseaux de neurones convolutifs, peut être une très bonne alternative à la longue et coûteuse procédure d'inversion mise en place par ces chercheurs pour l'estimation de la pression artérielle cérébrale.

Pour être en mesure de mener à bien ces différents objectifs et répondre efficacement aux différentes questions soulevées ci-dessus, l'ossature de ce manuscrit est organisée autour de sept chapitres repartis entre trois axes principaux comme suit :

Partie I - Concepts généraux et objectifs de la thèse

Composée de trois chapitres, cette partie introductive a pour but de fournir au lecteur les rudiments nécessaires à la compréhension de ce manuscrit de thèse.

Chapitre 1 - Introduction au système cardiovasculaire

De nature purement terminologique, descriptive et explicative, ce premier chapitre vise à fournir au lecteur et aux non-initiés les connaissances rudimentaires sur le système cardiovasculaire humain et les maladies y afférentes. La complexité du système cardiovasculaire y est juste esquissée afin de prioriser le but final de ce chapitre qui est d'introduire le lecteur à l'essentiel des termes clés liés au système cardiovasculaire humain et utilisés dans ce manuscrit.

Chapitre 2 - Bref aperçu des modèles cardiovasculaires actuels et des principaux outils mathématiques utilisés dans la thèse.

Ce deuxième chapitre a pour but d'introduire le lecteur aux modèles et outils mathématiques utilisés dans cette thèse. Pour ce faire, il présente tout d'abord les modèles physiques du système cardiovasculaire existants dans la littérature (i.e., les modèles couramment utilisés pour simuler et analyser le comportement physiologique et pathologique réel du système cardiovasculaire), tout en mettant l'accent sur le type de modèle utilisé dans ce travail, à savoir les modèles d'ordre réduit 0D. Ensuite, il introduit les rudiments des outils mathématiques utilisés dans cette thèse: le fonctionnement du filtre de Kalman d'ensemble précédemment utilisé dans la plateforme d'inversion pour estimer des paramètres hémodynamiques est rappelé,

tandis que les bases de l'analyse d'incertitude, de l'analyse de sensibilité et de l'apprentissage machine sont introduites.

Chapitre 3 - Contexte général, objectifs et plan de la thèse.

Après avoir introduit dans les deux chapitres précédents les concepts généraux utilisés dans cette thèse, le contexte général de la thèse ainsi que ses objectifs et contributions sont exposés dans ce dernier chapitre introductif. Le plan global de la thèse y est aussi donné ainsi que les principaux résultats de chaque chapitre.

Partie II - Analyse d'incertitude et de sensibilité

Constituée de deux chapitres, cette partie tend vers le premier objectif majeur de la thèse. D'une part, elle examine la précision et la robustesse de la procédure d'inversion proposée par les chercheurs, en lien avec diverses sources d'incertitude liées au modèles utilisés, aux hypothèses formulées et aux données cliniques du patient, et d'autre part, elle fixe un critère d'arrêt pour l'algorithme basé sur le filtre de Kalman d'ensemble utilisé dans la procédure d'inversion.

Chapitre 4 - Analyse de sensibilité rétrograde et covariance d'ordre réduit dans l'identification non invasive des paramètres des artères cérébrales.

Dans ce chapitre, une approche prenant en compte les incertitudes liées au traitement des images médicales du patient et aux hypothèses formulées sur les modèles utilisés, telles que les hypothèses liées aux conditions limites, aux paramètres physiques et physiologiques, est d'abord présentée pour quantifier les incertitudes sur les résultats de la procédure. Les incertitudes liées à la segmentation des images sont modélisées à l'aide d'une distribution gaussienne et celles liées au choix des hypothèses de modélisation sont analysées en testant plusieurs scénarios de choix d'hypothèses possibles. De cette démarche, il ressort que les incertitudes sur les résultats de la

procédure sont du même ordre de grandeur que celles liées aux erreurs de segmentation. Par ailleurs, cette analyse montre que les résultats de la procédure sont très sensibles aux hypothèses faites sur les conditions aux limites du modèle de flux sanguin. En particulier, le choix des conditions limites symétriques de Windkessel pour le modèle s'avère être le plus approprié pour le cas du patient étudié.

Enfin, ce chapitre propose également une construction déterministe pour la propagation rétrograde de l'incertitude permettant une estimation à faible complexité (tout aussi fiable que l'estimation par l'approche ensembliste) de la matrice de covariance des paramètres hémodynamiques du réseau artériel du patient après l'inversion.

Chapter 5 - Analyse de sensibilité globale pour évaluer l'importance des paramètres et fixer un critère d'arrêt dans un problème inverse biomédical.

Le chapitre 5 est en quelque sorte une extension de l'analyse d'incertitude et de sensibilité effectuée dans le chapitre précédent. Il consiste en une démarche permettant de classer les paramètres hémodynamiques estimés à l'aide de la procédure d'inversion par ordre d'importance et de fixer un critère d'arrêt pour l'algorithme utilisé dans cette procédure. Concrètement, il montre comment obtenir, en plus des écarts-types disponibles à l'issue de la procédure d'inversion, une répartition de l'incertitude totale dans les résultats du modèle de flux sanguin utilisé (les résultats de la procédure d'inversion) en petites portions d'incertitude dues à chaque paramètre d'entrée ou groupe de paramètres d'entrée du modèle. À cet effet, certains indicateurs statistiques pour l'analyse des paramètres d'un modèle, tels que les indices de Sobol, sont calculés. Les résultats de cette stratégie montrent, d'une part, que la plupart des résistances proximales sont les paramètres les plus importants pour l'estimation du débit sanguin dans les carotides internes et, d'autre part, que l'algorithme d'inversion peut être arrêté dès qu'un certain seuil de convergence raisonnable de ces paramètres

les plus influents est atteint. Ce critère d'arrêt a l'avantage de ne demander qu'un calcul supplémentaire très limité et d'éviter une sur-résolution inutile du problème; ce qui permet de réduire considérablement le temps de calcul de la procédure d'inversion.

Partie III - Apprentissage machine pour la pression cérébrale

Cette dernière partie du manuscrit a pour but d'atteindre le second objectif majeur visé par cette thèse. Constituée d'un seul chapitre, elle montre comment l'apprentissage machine, orienté réseaux de neurones convolutifs, peut être une très bonne alternative à la longue et coûteuse procédure d'inversion précédemment mise en place par Lal et al. [12] pour l'estimation de la pression artérielle cérébrale.

Chapitre 6- Imagerie médicale fonctionnelle d'un patient, augmentée par apprentissage implicite sur des variétés

Dans ce chapitre, une nouvelle plateforme numérique basée sur l'apprentissage machine permettant d'estimer la pression artérielle spécifique au patient dans les artères cérébrales beaucoup plus rapidement qu'avec la procédure d'inversion mais avec la même précision, est présentée. L'application de cette plateforme aux données du patient utilisées dans la procédure d'inversion permet une estimation non invasive et en temps réel de la pression cérébrale du patient cohérente avec l'estimation de la procédure d'inversion.

Conclusion

Chapitre 7- Résultats principaux, discussions et perspectives

Dans ce dernier chapitre, les conclusions et les principaux résultats tirés de cette thèse sont rappelés. Ensuite, les différents choix de modélisation physique, les hypothèses

formulées ainsi que leurs conséquences sont discutés. Enfin, de possibles orientations et perspectives de recherche future sont proposées.

Contents

List of Abbreviations

AA	**A**scending **A**orta
ANNs	**A**rtificial **N**eural **N**etwork**s**
BAF	**B**rain **A**neurysm **F**oundation
CNNs	**C**onvolutional **N**eural **N**etwork**s**
CHRU	**C**entre **H**ospitalier **R**égional **U**niversitaire
coW	**c**ircle **o**f **W**illis
CV	**C**oefficient of **V**ariation
CVDs	**C**ardio**v**ascular **D**isease**s**
CVS	**C**ardio**v**ascular **S**ystem
DBP	**D**iastolic **B**lood **P**ressure
EnKF	**E**nsemble **K**alman **F**ilter
GSA	**G**lobal **S**ensivity **A**nalysis
ICAs	**I**nternal **C**arotid **A**rtorie**s**
KF	**K**alman **F**ilter
L-ICA	**L**eft **I**nternal **C**arotid **A**rtery
LSA	**L**ocal **S**ensivity **A**nalysis
MAE	**M**ean **A**bsolute **E**rror
ML	**M**achine **L**earning
MRA	**M**agnetic **R**esonance **A**ngiogram
MRI	**M**agnetic **R**esonance **I**maging
NNs	**N**eural **N**etwork**s**
PC-MRI	**P**hase **C**ontrast **M**agnetic **R**esonance **I**maging
R-ICA	**R**ight **I**nternal **C**arotid **A**rtery
RSDs	**R**elative **S**tandard **D**eviation**s**
SA	**S**ensivity **A**nalysis
SBP	**S**ystolic **B**lood **P**ressure
SDs	**S**tandard **D**eviation**s**
TL	**T**ransfer **L**earning
UA	**U**ncertainty **A**nalysis
UQ	**U**ncertainty **Q**uantification
WHF	**W**orld **H**eart **F**ederation
WHO	**W**orld **H**ealth **O**rganization
WK3	**W**indkessel **T**hree element

List of Symbols

α	First parameter used to model the patient-specific inlet flow rate.
ϵ	Third and last parameter used to model the patient-specific inlet flow rate.
μ	Blood viscosity.
ρ	Blood density.
τ	Second parameter used to model the patient-specific inlet flow rate.
a_C	Multiplicative constant used to introduce the nonsymmetry in the Windkessel compliances.
a_R	Multiplicative constant used to introduce the nonsymmetry in the Windkessel resistances.
C	Arterial compliance.
D	Vascular distensibility.
E	Young modulus.
G	Vascular conductance.
h	Arterial wall thickness.
k_1	First constant used to model Young modulus.
k_2	Second constant used to model Young modulus.
k_3	Third and last constant used to model Young modulus.
L	Blood inertia.
l	Arterial length.
R	Vascular resistance.
r	Arterial radius.
R_D	Windkessel distal resistance.
R_P	Windkessel proximal resistance.
V	Vascular volume.

List of Figures

List of Tables

Part I

General concepts and objectives

CHAPTER 1

An introduction to the cardiovascular system

Chapter contents

Before getting to the heart of the matter, an introduction to the cardiovascular system is first proposed to the reader. Purely terminological, descriptive and explanatory in nature, this chapter provides a comprehensive review of the necessary knowledge about the cardiovascular system for the uninitiated and carefully introduces the bulk of the key terms subsequently used in this manuscript. The complexity of the cardiovascular system is just sketched here. The interested reader can get a more detailed presentation by consulting a relevant and dedicated bibliography as [13, 14, 15, 16, 17].

1.1 Some basic anatomical and physiological aspects of the cardiovascular system

The human body is composed of a number of systems that control different functions, such as the respiratory system for breathing, gas exchange, phonation and olfaction; the digestive system for the physical and chemical breaking down of ingested food and the elimination of non-digestible wastes; and the nervous system (major controlling, regulatory, and communicating system in the body) for all mental activity including thought, learning, reacting to the environment, and memory. This Section focuses on the cardiovascular system (CVS), which is the system responsible for conveying blood to and from all parts of the body.

1.1.1 Structure of the cardiovascular system

Like other systems in the body, the CVS, also called the circulatory system or the vascular system, is made up of three major organs: the heart, blood, and blood vessels (i.e., arteries, veins, arterioles, venules, and capillaries) [18, 19]. Together these organs

carry out the complex activities of the CVS. The function of the heart is to act as the pump, pushing blood into the blood vessels. The blood vessels carry blood to all the different parts of the body, before it is returned back to the heart to start the cycle again. This process of blood circulation, which includes the intake of metabolic products, the distribution of these products throughout the body, and the return of harmful by-products (carbon dioxide, water, heat, etc.) to the environment [20, 21], is known as the circulatory loop.

1.1.1.1 The circulatory loop

The circulatory loop mainly consists of two components: the systemic circulation and the pulmonary circulation [22, 23, 24].

The pulmonary circulation is the portion of the CVS which transports oxygen-depleted (de-oxygenated) blood (parts in blue in Figure 1.1) from the right side of the heart, via the pulmonary artery, to the lungs. Gas exchange occurs in the lungs, whereby carbon dioxide is released from the blood, and oxygen is absorbed [19]. The now oxygen-rich blood is then returned to the left ventricle via the pulmonary vein [25, 26, 27].

The systemic circulation, also called the greater circulation or peripheral circulation, is the portion of the CVS which conveys highly oxygenated blood (parts in red in Figure 1.1) away from the heart through the aorta from the left ventricle where the blood has been previously deposited from pulmonary circulation, to the rest of the body (with the exception of the lungs and the heart itself, which is supplied with oxygen and nutrients through a small "loop" of the systemic circulation, called the coronary circulation). During this process, wastes from body tissues are removed and oxygen-depleted (deoxygenated) blood is returned back to the right ventricle [25, 26, 27].

Thus, this closed tubular system that is the CVS can be physiologically thought as the transport, protection and regulation system of the body: it is responsible for transporting nutrients (such as amino acids and electrolytes), respiratory gases (oxygen,

carbon dioxide), hormones, metabolic materials and blood cells throughout the organism to provide nourishment, to help in fighting diseases, to stabilize temperature and pH, to maintain homeostasis [26, 28, 29, 30], and in general, to maintain an appropriate environment in all the tissue fluids of the body for optimal survival and function of the cells [31, 32, 33]. All this is achieved by means of blood, which is constantly pumped by the heart, and circulates through a very complex network composed of capillaries, veins and arteries.

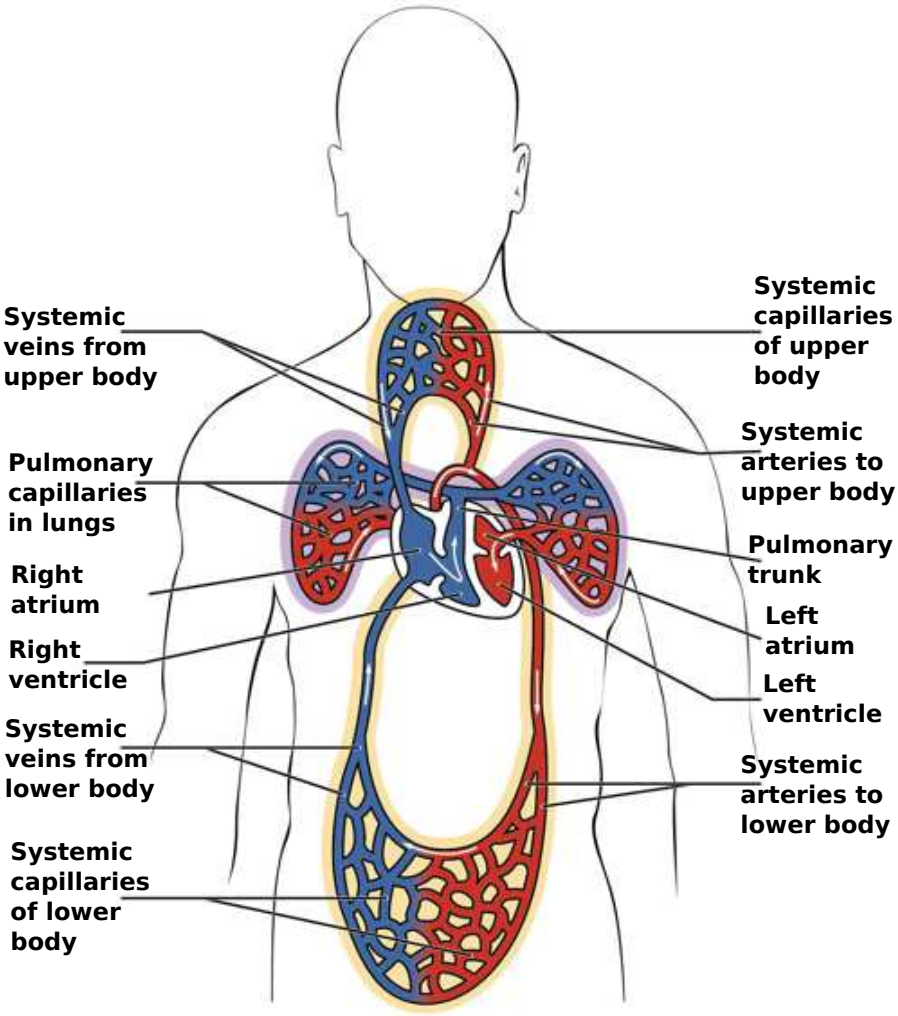

Figure 1.1: Sketch of the distribution of blood in the different parts of the circulatory system. Adapted from [13].

1.1.1.2 The arteries

1.1.1.2.1 Basic structure of an artery

As shwon in Figure 1.2, in general, an artery (no matter what its size) mainly consists of the following three layers [34, 35, 36, 37]:

1. The **tunica adventitia** or **tunica externa** is the outermost layer. From Latin words literally meaning "coat that comes first", referring to how it is found during the dissection of a vessel, this layer provides structural support and shape to the artery. It is made of strong, flexible fibrous connective tissue (collagen fibers and elastic tissue). Its main function is to help hold arteries open and prevents tearing of the arterial walls during body movements.

2. The **tunica media**, or **media**, is the middle layer. From Latin words meaning "middle coat", this layer is made up of smooth muscle cells, collagen fibres, sandwiched together with elastic connective tissue. According to some anatomists, its elastic portion is distinct enough to call it a separate elastic layer of the wall. The smooth muscles surrounding the tunica media permit to regulate the internal diameter of the artery.

3. The **tunica intima**, or simply the **intima**, which is in direct contact with the flow of blood, is the innermost layer. From Latin words translating to mean "inner coat", this layer is made up of an endothelial lining which provides a frictionless pathway for movement of blood and a supporting layer of elastin rich collagen in elastic arteries. The hollow internal cavity in which the blood flows is called the lumen.

1.1.1.2.2 Classification of arteries

On the basis of their structure, their functions, and particularly according to the relative composition of elastic and muscle tissues in their tunica media, the arteries of the systemic circulation can be subdivided into muscular or elastic types. Larger arteries

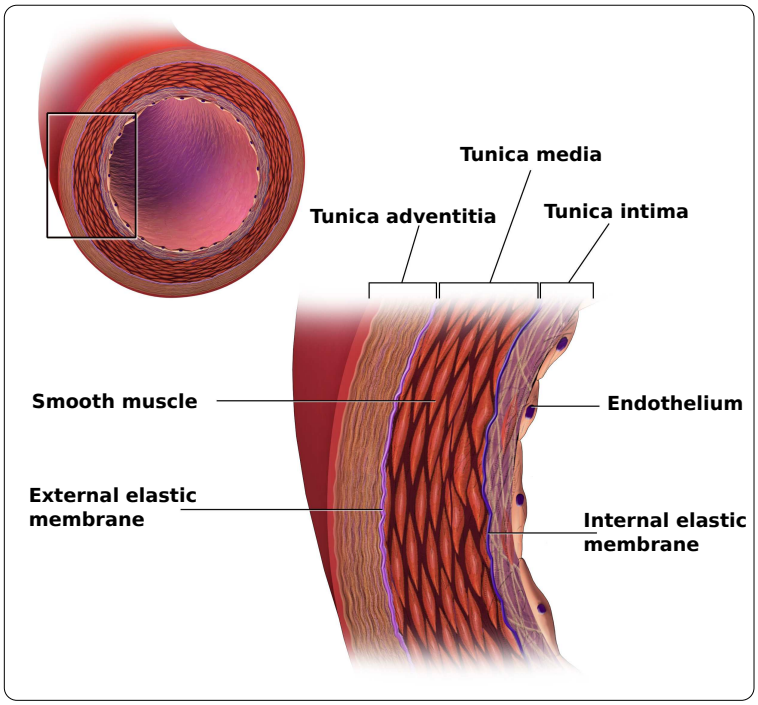

Figure 1.2: Structure of an artery wall. Adapted from [38].

are typically elastic and smaller arteries are more likely to be muscular [39]. These arteries deliver blood to the arterioles, which in turn deliver blood to the capillary networks associated with the body's tissues [39]. Thus, in general, the arteries are classified into the following three main categories [39, 40]:

- **Elastic arteries**, also called conducting arteries or conduit arteries, are arteries with a large number of collagen and elastin filaments in their tunica media, giving them the ability to stretch in response to each pulse [39, 41]. This feature of being elastic gives rise to the Windkessel effect, which allows arteries to maintain a relatively constant pressure gradient despite the pulsating nature of the blood flow from the heart [39, 42]. The largest arteries in the body, those nearest the heart, such as the aorta, common carotid, subclavian, pulmonary and common iliac arteries comprise the body's system of elastic arteries [39]. They are connected to muscular arteries.

- **Muscular arteries**, also called distributing arteries, are medium-sized arteries

that draw blood from elastic arteries and branch into resistance vessels, including small arteries and arterioles [39, 43]. They include the anatomically named arteries like the brachial artery, the radial artery, the femoral artery, and the splenic artery, to cite a few [42]. Their basic function is to distribute blood to the different organs of the body.

- **Arterioles** are small diameter blood vessels in the microcirculation system that branch out from arteries and lead to capillaries [39, 44]. They provide blood to the organs and are mostly composed of smooth muscle [42, 45]. They play a significant role in the systemic vascular resistance because of the lack of significant elastic tissue in the walls [42]. This feature of vascular resistance reduces the pressure and velocity of blood flow in arterioles to enable gas and nutrient exchange to occur within the capillaries [39].

1.1.1.2.3 The circle of Willis

The greater part of the human brain is supplied with arterial blood by an arrangement of interconnected arterial channels in the form of a polygon (roughly considered as an heptagon or a circle) called circle of Willis (coW) or circulus arteriosus, Willis' circle, loop of Willis, cerebral arterial circle, or Willis polygon [46, 47, 48]. The coW is therefore an arterial ring located just at the base of the brain (see Figure 1.3) [49, 50]. It is named after Thomas Willis (1621-1675), an English physician [50, 51]. Well-known as the most common site of cerebral aneurysms [52, 53], its anatomy and its physiology are well described in [54] as follows:

The anastomosis of two major sets of arteries, namely the internal carotid arteries (left and right) and the vertebral arteries (left and right) contributes to its formation [55]. The internal carotid arteries supply blood to the greater part of the brain. They also have branches that supply the face and scalp. Anteriorly, the two anterior cerebral arteries arise from the internal carotid arteries and are joined by the anterior communicating artery. Posteriorly, the two vertebral arteries join to form the

basilar artery near the brain stem at the base of the skull. After travelling for a short distance the basilar artery divides to form two posterior cerebral arteries, each of which is joined to the corresponding internal carotid artery by a posterior communicating artery, completing the circle. From this circle, the anterior cerebral arteries pass forward to supply the anterior part of the brain, the middle cerebral arteries pass laterally to supply the sides of the brain, and the posterior cerebral arteries supply the posterior part of the brain. Branches of the basilar artery supply parts of the brain stem. Thus, the coW, as shown in Figure 1.3 is formed by: two anterior cerebral arteries, two internal carotid arteries, one anterior communicating artery, two posterior communicating arteries, two posterior cerebral arteries, and one basilar artery [56].

The arrangement of these arteries in the coW is generally recognized – although this theory is no longer universally accepted nowadays (see for example [50]) – as an example of collateral circulation, a compensatory system in the case of arterial occlusion or stenosis [57, 58, 59, 60, 61, 62, 63, 64, 65]. This means that the coW is such that the brain as a whole receives an adequate blood supply even when a contributing artery is damaged or obstructed or during extreme movements of head and neck [46, 55].

1.1.2 Blood flow, blood pressure, resistance, conductance and their interrelationships

In this Section, the excellent description of the interrelationships among blood flow, pressure, resistance, and conductance as presented in [32] is followed.

1.1.2.1 Blood flow

Often expressed in ml/min or in L/min but also in ml/s or in any other unit of flow, **blood flow**, Q, simply refers to the quantity of blood that passes a given point in the circulation in a given period of time.

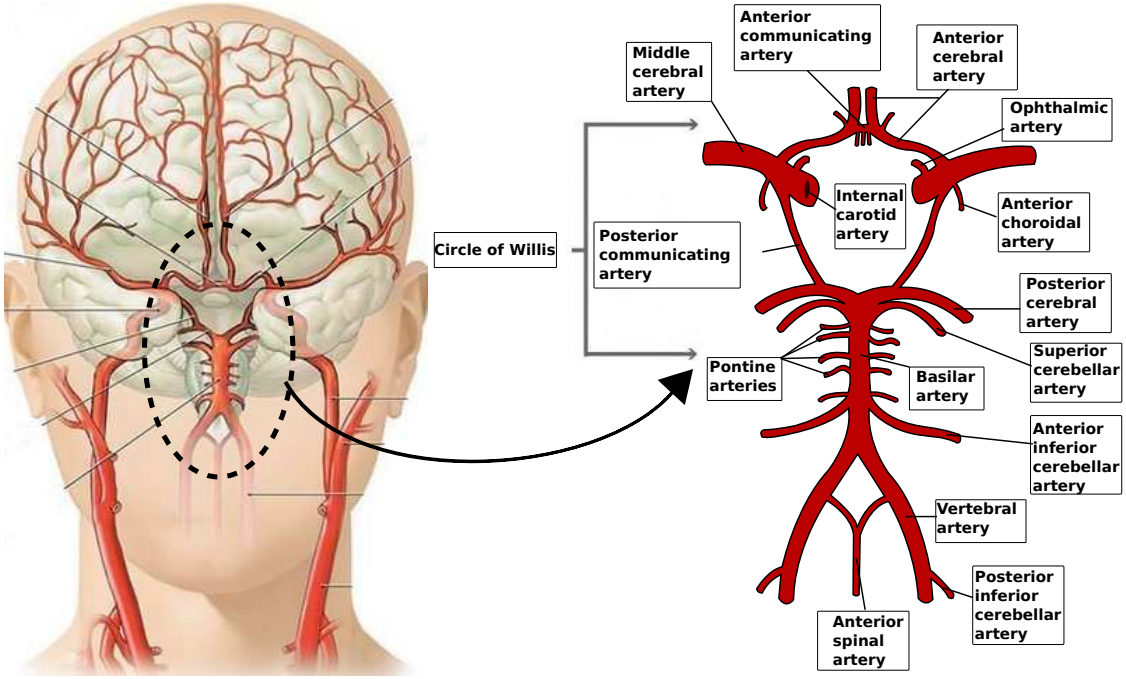

Figure 1.3: Anterior view of head showing the major arteries of the brain. The enlargement of the encircled area shows the arteries of the brain and brain stem, including a complete structure of the circle of Willis. Adapted from [66, 67].

1.1.2.2 Blood pressure

The term **blood pressure**, P, usually measured in millimeters of mercury (mmHg) refers to the force exerted by the blood against any unit area of the vessel wall. Nowadays, there are mainly two methods for measuring blood pressure [19]: the one using sphygmomanometer (a device with a cuff that, when placed around the arm, inflates for about a minute providing a pressure measure at a single point in time), and that using the catheter (a device commonly placed in the radial or femoral artery), for a continuous pressure measure. When blood pressure is measured using a cuff, it provides two meaningful values: the higher one is the Systolic Blood Pressure (SBP), observed during ventricular contraction, and the lower one, reached during ventricular relaxation (diastole), is Diastolic Blood Pressure (DBP) [68].

Blood pressure is also well known as one of the key information in the medical field. Indeed, the arterial blood pressure and how it varies over time are considered to be a key

indicator of cardiovascular health, including several clinical conditions [69, 70, 71, 72]. A high level of blood pressure may indicate that there is a risk of more severe cardiovascular diseases. Also, in a hospital context, low pressure values in very sick patients might indicate blood loss and other critical situations. Therefore, there is much interest in developing a reliable noninvasive technique for blood pressure assessment, in particular in the arteries where occur cardiovascular diseases.

1.1.2.3 Vascular resistance and conductance

Vascular resistance, R, represents the impediment, the hindrance to blood flow in a vessel at a given pressure difference. This cannot be measured by any direct means. It is calculated from measurements of blood flow and pressure difference between two points in the vessel and is often related to the notion of **vascular conductance**, G, (generally expressed in ml/s/mmHg), which is the ease with which the blood flows through a vessel for a given pressure difference. Therefore, vascular resistance is the exact reciprocal of conductance in accord with the following equation [73]:

$$G = \frac{1}{R} \tag{1.1}$$

where G and R are vascular conductance and resistance, respectively.

1.1.2.4 Interrelationships among blood blow, pressure, and resistance

As shown in Figure 1.4, blood flow through a vessel is determined by two factors: the blood pressure difference between the two ends of the vessel, also known as "pressure gradient" along the vessel, which is the force that pushes the blood through the vessel, and the vascular resistance. Therefore, blood flow through a vessel can be calculated by the following formula:

$$Q = \frac{\Delta P}{R} \tag{1.2}$$

where Q, ΔP, and R are blood flow, the pressure difference ($P_{in} - P_{out}$) between the two ends of the vessel (see Figure 1.4), and the vascular resistance, respectively. This

formula states, in effect, that the blood flow is directly proportional to the pressure difference but inversely proportional to the resistance. However, such a relationship is strictly valid only when flow does not vary over time. With a time-varying flow and under a broad range of conditions, vascular resistance is approximately independent of the flow rate, and depends only on the geometrical properties of the vessel and on the blood viscosity [74].

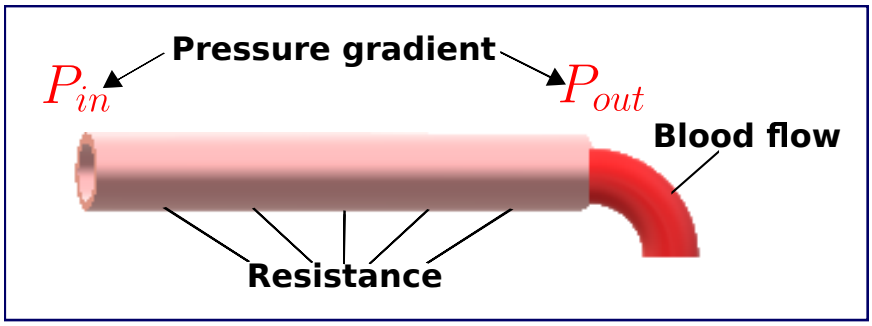

Figure 1.4: A blood vessel segment located anywhere in the circulatory system showing the interrelationships among pressure, flow, and resistance. P_{in} represents the pressure at the origin of the vessel; at the other end, the pressure is P_{out}. Resistance occurs as a result of friction between the flowing blood and the intravascular endothelium all along the inside of the vessel. Adapted from [32, 75].

1.1.3 Vascular distensibility and compliance

Distensibility and compliance are important vessel wall properties. Distensibility is related to elastic properties of the arterial wall, and compliance reflects the buffering function of the artery [76]. Below the excellent description of these vessel characteristics detailed in [32] is followed

1.1.3.1 Vascular distensibility

One valuable property of the CVS is that all blood vessels are distensible [32, 75]. This is visible when blood pressure increases. This dilates the blood vessels and therefore decreases their resistance. As a result, blood flow increases not only because of increased pressure but also because of decreased resistance [32, 75]. Vascular distensibility is

also noticeable in circulatory function. Indeed, because of their distensible nature, the arteries are able to accommodate the pulsatile output of the heart and to average out the pressure pulsations [75]. This provides smooth, continuous flow of blood through the very small blood vessels [32]. Therefore, **vascular distensibility**, D, is commonly defined as the relative change in vascular volume, V, per unit of pressure, i.e., fractional increase in volume for each millimeter of mercury rise in pressure, in accordance with the following formula [32, 75]:

$$D = \frac{\frac{\Delta V}{V}}{\Delta P} = \frac{\Delta V}{V \times \Delta P} \tag{1.3}$$

where V, ΔV, and ΔP are original vascular volume, increase in vascular volume, and increase in vascular pressure, respectively.

1.1.3.2 Vascular compliance

In hemodynamic studies, another feature of the vascular system, much more important to know than the distensibilities of the individual vessels, is **vascular compliance** [32]. Also called vascular capacitance, it refers to the total quantity of blood that can be stored in a given portion of the circulation for each millimeter of mercury rise in pressure [32]. In other words, vascular compliance is the change in vascular volume per unit of pressure [77]. Thus, although vascular compliance, C, and vascular distensibility, D, are quite different [32], using Equation (1.3), they can be related by this formula:

$$C = \frac{\Delta V}{\Delta P} = D \times V \tag{1.4}$$

where C, D, V, ΔV, and ΔP are vascular compliance, vascular distensibility, original vascular volume, increase in vascular volume, and increase in vascular pressure, respectively.

1.1.4 Blood inertia

Blood inertia refers to the ability of blood to prevent sudden blood movement at each heart beat. In other words, blood inertia represents the resistance of blood to any

change in its flow velocity at each heart beat; it is of course related to the density of blood, ρ.

Ultimately, all these various characteristics of the vascular system presented above play an important role in maintaining a healthy CVS. In consequence, their dysfunction might increase the risk of arterial wall damage, which could inevitably give rise to cardiovascular diseases.

1.2 Cardiovascular diseases

Cardiovascular diseases (CVDs) are a group of disorders affecting the structure or function of the heart and blood vessels. These include coronary heart disease, arrhythmias (abnormal heart rhythms), cerebrovascular disease, peripheral arterial disease, rheumatic heart disease, congenital heart disease, deep vein thrombosis, pulmonary embolism, and so on [78]. According to the WHO statistics, CVDs are the number one cause of death globally [78]: more people die annually from CVDs than from any other cause and in 2016, 17.9 million people died due to CVDs, representing 31% of all global deaths. Out of these deaths, 85% are due to heart attack and stroke.

These statistics challenge the research community, and through it the international community as a whole, on the need not to take this global scourge lightly. However the exact cause of CVDs is not clear [79], even if several well-known risk factors are associated with them. Some risk factors, such as family history, age, gender, ethnicity, socioeconomic status, co-existing diseases (diabetes or kidney disease, for example), cannot be modified, while other risk factors, which are much more behavioural, such as unhealthy/unbalanced diet, physical inactivity, tobacco use and harmful use or excessive consumption of alcohol, are modifiable [80]. The effects of the latter may show up in individuals as “intermediate risks factors” such as, hypertension (high blood pressure), hyperlipidaemia (high cholesterol, high blood lipids), diabetes (high blood glucose), and overweight and obesity. These indicate an increased risk of developing a heart attack, stroke, heart failure and other complications. Of all these risk factors, hypertension is

the single biggest risk factor of CVDs, which is widely used to diagnose heart attack or stroke [80]. Indeed, according to the World Heart Federation (WHF) statistics, at least 970 million people worldwide are suffering from hypertension and number is increasing day by day so that it is estimated that in twenty-five years, 1.56 billion adults will suffer from high blood pressure [81]. Therefore, there is a real need for a technique to noninvasively measure the pressure in any artery in the body.

1.2.1 Atherosclerosis

As shown in Figure 1.5, **atherosclerosis** is a form of CVD involving the hardening and the narrowing of an artery specifically due to a build up of an atheromatous plaque (usually made up of fatty substances, cholesterol, cellular waste products, calcium and fibrin), in the inner lining of the medium-sized or large arteries (not just there, but often) [82]. In atherosclerosis, the plaques cause the arteries to harden and narrow, limiting the flow of oxygen-rich blood to vital organs, and increasing the risk of blood clots that could potentially block the arteries [83, 84, 85]. This leads to failure of the organs and structures, ischemia, a restriction in blood supply to organs or structures tissues, or other serious health risks, including angina or heart attack, stroke, or even death [83, 86]. In general, this disease is one of the major cause of high blood pressure.

1.2.2 Cerebral aneurysm

A **cerebral aneurysm** (also called intracranial aneurysm or brain aneurysm) is a bulging area, an abnormal dilation (widening or balloning) in the wall of an artery in the brain, analogous to a thin balloon or a weak spot on a tire's inner tube [88, 89]. This is caused by an abnormal loss, absence, or abnormal degenerative (breaking down) change (or weakening) of the inner layer in the artery and also by the effects of pressure from the pulsations of blood pumped forward through the arteries in the brain [88]. Cerebral aneurysms commonly tend to occur at the junctions between the arteries making up the

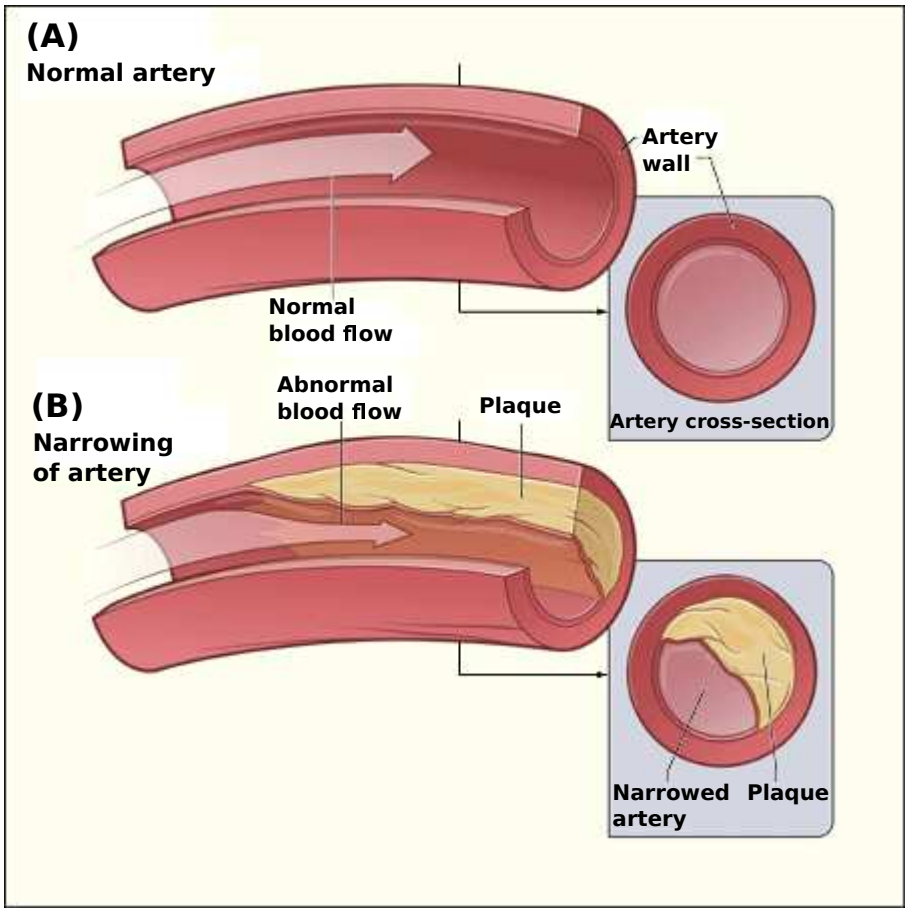

Figure 1.5: Atherosclerosis: a normal artery with normal blood flow together (A). An artery with plaque buildup (B). The inset images show a cross-section of a normal artery (A) and a cross-section of an artery with plaque buildup (B). Adapted fom [87].

coW [52, 53, 88]. However, arteries anywhere in the brain can develop aneurysms [88].

In general, cerebral aneurysms, be they saccular (berry), fusiform or dissecting, or be they small, large, giant or super-giant, are known as silent killers, because in most of the cases they have no symptoms and are typically found incidentally when tests are being done to screen for other conditions [88, 89]. Because their walls may be weak and thin, aneurysms are at risk of rupturing. But a ruptured aneurysm can cause life-threatening blood loss, which leads to death [86]. Indeed, when an aneurysm ruptures, blood spills into the space between the skull and the brain, causing subarachnoid hemorrhage (SAH) [89].

Many risk factors can contribute to the formation and rupture of cerebral aneurysms. One of the most significant factors is high blood pressure (hypertension), which can fortunately be controlled. There is therefore great interest in developing techniques and CVS models allowing noninvasive and reliable blood pressure prediction.

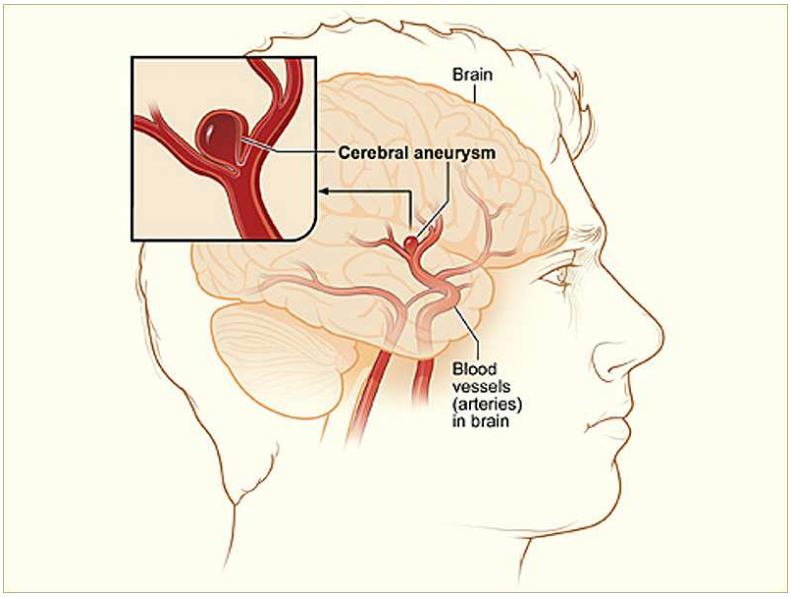

Figure 1.6: Cerebral aneurysm. Taken from [90].

Chapter key points:

- **The CVS is a complex system of interconnected organs whose dysfunction is the source of many diseases.**
- **Knowledge of the cardiovascular anatomy and of interrelationships among blood flow, pressure and resistance is a prerequisite for understanding how the CVS works.**
- **Arteries, especially cerebral arteries and associated diseases such as cerebral aneurysms are the region of interest in this work.**

CHAPTER 2

A brief overview of current cardiovascular models and main mathematical tools

Chapter contents

This second chapter of the first part introduces the reader to the models and mathematical tools used in this thesis. It begins with a literature review on existing physical models of the CVS (i.e., the models commonly used to simulate and analyze actual physiological and pathological behavior of the CVS), with emphasis on the type of model used in this work, namely the

lumped-parameter models. This literature review is essentially based on the work of Shi et al. [91], Kokalari et al. [92], and Zhou et al. [93]. Then, it introduces the basics of the mathematical tools used in this thesis: the ensemble Kalman filter technique previously used in the inversion platform to estimate hemodynamic parameters [12] is recalled, while the basics of uncertainty analysis, sensitivity analysis and machine learning practices are introduced.

2.1 Types of existing physical models of the CVS

The mathematical description and physical modeling of the human CVS have been proven nowadays to be very useful prerequisite tools for understanding the complex physiological processes and mechanisms of a real system by simplifying it. They have become powerful tools for simulating the hemodynamic properties of CVS and have been playing an increasingly important role in the diagnosis or the treatment of CVDs and in the development of medical devices [94, 95, 96, 97, 98]. That is why, with the recent significant improvements in computer technology, modeling based on physical principles has attracted increasing interest.

At present, as illustrated in Figure 2.1, the types of physical models of the CVS existing in the literature include two types of models: high-dimensional models and low-dimensional models [92, 93].

2.1.1 High-dimensional models

In general, high-dimensional modeling is only dedicated to detailed descriptions of the local complex hemodynamics of a specific region in the CVS (see for example the work carried out by the YALES2BIO[1] biomechanics team, named after the YALES2BIO solver used for computational fluid dynamics [27, 99, 100, 101, 102, 103, 104, 105, 106, 107]). However, the implementation of a high-dimensional model for the whole

[1] https://imag.umontpellier.fr/~yales2bio

arterial system is computationally impracticable due to the enormous complex physical, geometrical, mechanical and numerical parameters requested. The well-known high-dimensional models of the CVS available in the literature to date mainly include **2D models** and **3D models**. While 2D models are usually used to describe changes in the radial velocity of blood flow in an axisymmetric tube [108, 109], 3D models are generally applied to simulate the fluid-structure interaction between the vascular walls and blood [110, 111].

2.1.2 Low-dimensional models

Unlike high-dimensional modeling, the low-dimensional modeling, because of its small computational cost can be used in practice to describe the hemodynamic characteristics of a patient-specific entire arterial tree. This explains why the low-dimensional models have been extensively used in the study of cardiovascular dynamics. The three main types of low-dimensional models commonly used in the literature are **0D models**, **1D models**, and **tube-loads models** [92, 93].

While 0D models are generally used to describe global properties of the arterial system, 1D models, commonly based on the simplified Navier-Stokes equations, are used to simulate pressure and flow waveforms at any point of the whole arterial network according to their distributed properties [112, 113, 114, 115, 116]. As for tube-loads models, which are developed by taking advantage of simplicity of 0D models and accuracy of 1D models [117, 118], they are most often used for monitoring several hemodynamic parameters such as pulse transit time, arterial compliance, pulse wave velocity, etc. [93].

According to the pulse waveforms spatial distribution within them, low-dimensional and high-dimensional models are usually grouped into two categories: the **lumped-parameter models** (0D models), characterized by pulse waveforms which are only functions of time, and the **distributed models** (1D models, tube-load models together with high dimensional models), characterized by pulse waveforms

which are both functions of time and space (see Figure 2.1). Of these models, the simplest and least computationally expensive one capable to describe global properties of the arterial system is 0D model. A more detailed description of this model is given below in next Section because of its use in this manuscript.

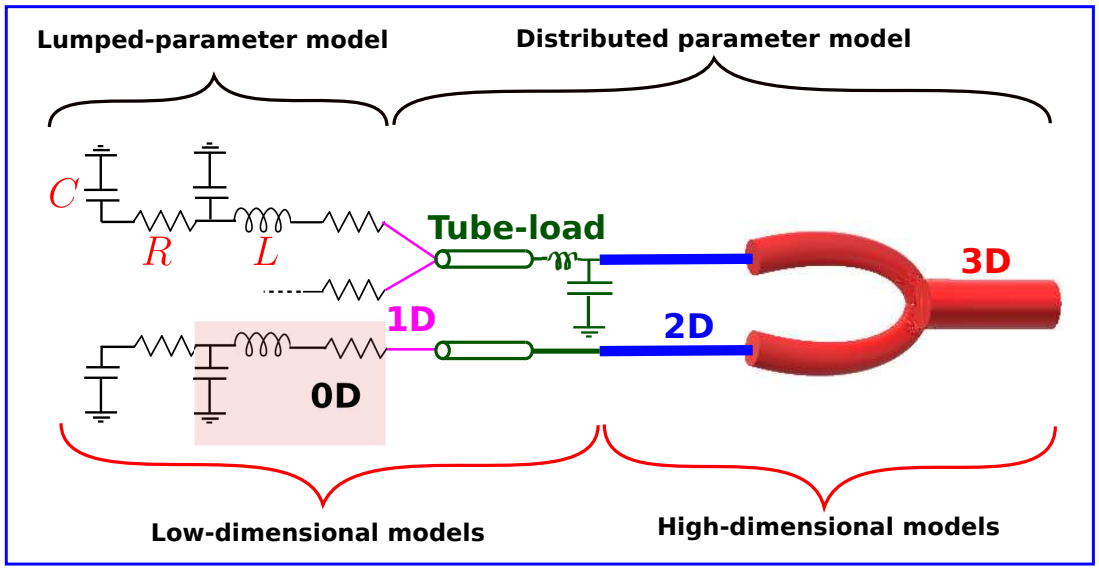

Figure 2.1: Schematic representation of physical models of the CVS. Adapted from [92, 93].

2.2 The lumped-parameter models of the CVS

Also called lumped-element models, lumped-component models, lumped-compartment models or lumped-parameter models, the 0D models have been developed and used by many researchers all over the world to mimic in a simple way the behavior of blood flow and pressure in the CVS [86, 112, 119, 120, 121, 122, 123, 124, 125, 126].

As illustrated in Figure 2.1, whilst the models of higher dimensions account for the spatial variation of fundamental variables (pressure, volume and flow rate), the 0D models assume as uniform the distribution of these variables within every single compartment (organ, blood vessel, part of blood vessel, etc.) of the model at any instant in time [91, 127]. The 0D models are therefore appropriate for studying the

global distribution of the pressure, flow rate, and blood volume over a range of specific physiological conditions and also for simulating the global hemodynamics of the CVS [91, 127].

In 0D modeling is often applied the analogy between the blood flow in the systemic arteries and the current in an electric circuit. Blood pressure gradient in arterial tree pushes the blood against the **hydraulic impedance** (i.e., combined effect of the total loss due to friction, the vessels capacitance due to their elasticity, and blood inertia) like voltage gradient in an electric circuit pushes current against the **electric impedance** (i.e., the combination of the resistance, capacitance, and inductance in the circuit) [91, 92]. Blood flow is described by the continuity equation for mass conservation, Poiseuille's law for the steady state momentum equilibrium and by the Navier-Stokes equation for the unsteady state momentum balance [91]. Analogically, current in an electric circuit is described by Kirchhoff's current law for current balance, Ohm's law for the steady state voltage-current relation and by the line transmission equation for the voltage-current relation in high frequencies [91]. Hence, as shown in Table 2.1, the effects of friction due to viscosity, the inertia in blood flow, the capacitance of the vessel, the blood pressure and the flow rate, can be respectively described by the resistance R, inductance L, capacitance C, voltage U, and the current I, in an electric circuit [92]. As consequence, the well-established methods used for studying the electric circuits can be applied to the investigation of cardiovascular dynamics too [92]. In doing so, each vascular compartment included in the lumped-parameter models simply boils down to two ordinary differential equations complemented by an algebraic balance equation that relates compartment volume to pressure [92]. The most well-known lumped-parameter model is the Windkessel model [128], which can divided into two subgroups: mono-compartment models and multi-compartment models [91, 127].

Table 2.1: Analogy between fluid dynamics and electric circuit. Adapted from [92]

Fluid dynamics	**Physiological variables**	**Electrical analog**
Pressure P	Blood pressure P	Voltage U
Flow rate Q	Blood flow rate Q	Current I
Volume V	Blood volume V	Charge q
Viscosity μ	Blood resistance $R = \frac{8\mu l}{\pi r^4}$	Electrical resistance R
Elastic coefficient	Vessel's wall compliance C	Capacitor's capacitance C
Inertance	Blood inertia L	Inductor's inertance L
Poiseuille's law	$Q = \frac{\Delta P}{R} = \frac{\Delta P \pi r^4}{8\mu l}$	Ohm's law $I = \frac{\Delta U}{R}$

2.2.1 Mono-compartment models

In mono-compartment models, the whole vessel network is treated as a whole — meaning that the internal pressure distribution and flow velocity in different parts of the vascular network are not calculated separately — and is represented with a single combination of the following elements: resistance, inductance and compliance (RLC). According to the number of these elements within them, current mono-compartment models are commonly classified into four main types [93]: two-element, three-element, four-element and complex mono-compartment Windkessel models. The details of all these models are beyond the scope of this manuscript. Only the two-element and three-element models are discussed below because of the use of three-element model in this work.

2.2.1.1 Two-element Windkessel model

First introduced by Stephen Hales in 1733 [129, 130, 131], and later formulated mathematically by Otto Frank in 1899 [128], the two-element Windkessel model is the precursor of mono-compartment models. Composed of a resistor, R, connected in parallel with a capacitor, C (see Figure 2.2A), the two-element was developed to

represent the elementary characteristics of the systemic artery network, while the veins were neglected and represented as a far field zero pressure. Applying Ohm's and Kirchhoff's laws of the electric circuits to the Windkessel model shown in Figure 2.2A, the governing differential equation describing the relationship between blood pressure and flow within the two-element Windkessel model can be obtained as follows:

$$q(t) = \frac{P(t)}{R} + C\frac{\mathrm{d}P(t)}{\mathrm{d}t} \tag{2.1}$$

where $q(t)$, $P(t)$, R, and C are blood flow rate, blood pressure, total peripheral resistance, and total arterial compliance, respectively. Since the development of this two-element model, many adaptations and different approaches have been proposed to describe the CVS.

2.2.1.2 Three-element model

Also called the Westkessel model, or RCR model, the three-element model, as shown in Figure 2.2B, is an extension of the Windkessel RC model by introducing an extra resistance element to improve the high frequency performance of the model [132]. Landes [133], and later Westerhof et al. [132], then Burattini and Natalucci [119], were the protagonists of this model even though each used it in a different configuration and for different purposes. Despite its shortcomings pointed out by *in vivo* studies [91, 134], the RCR model was widely used in cardiovascular simulations as the after-load for assessing cardiac function under various physiological and pathological conditions [91]. This model was also useful for describing the arterial characteristics, for capturing grossly the wave reflection response and frequency characteristics in Westerhof et al. [132], and for providing a good representation of after-load in the context of prediction of stroke volume, stroke work, and systolic and diastolic aortic pressure [134]. The governing differential equation for the RCR model can be obtained by applying Ohm's and Kirchhoff's laws of the electric circuits to the

three-element Windkessel model shown in Figure 2.2B. It can be expressed as follows:

$$\left(1+\frac{R_1}{R_2}\right)q(t)+CR_1\frac{\mathrm{d}q(t)}{\mathrm{d}t}=\frac{P(t)}{R_2}+C\frac{\mathrm{d}P(t)}{\mathrm{d}t} \tag{2.2}$$

where $q(t)$, $P(t)$, R_1, R_2, and C denote blood flow rate, blood pressure, the characteristic resistance (assumed to be equivalent to the characteristic impedance), the peripheral resistance, and vascular compliance, respectively. Compared with the two-element model, the three-element model can have a higher accuracy. Consequently, it has been extensively used in theoretical research [135, 136, 137]. A variant of the RCR model is the RLC model (see Figure 2.2C) used in this work. The latter model is commonly used as a building block in the development of multi-compartment models discussed in the next Section.

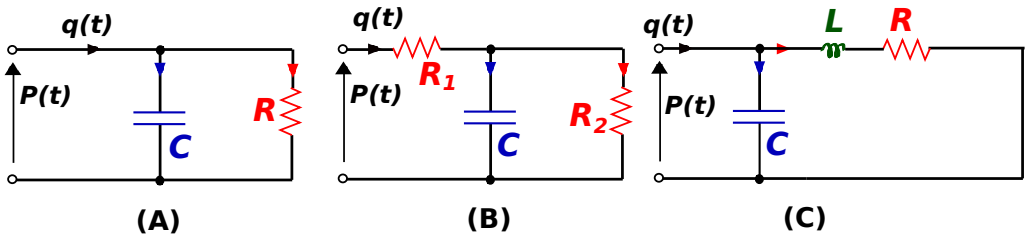

Figure 2.2: Mono-compartment models: A: Two-element Windkessel model. B: Three-element Windkessel model with two resistors and a capacitor. C: Three-element Windkessel model with a resistor, an inductor, and a capacitor.

2.2.2 Multi-compartment models

In the mono-compartment models explored above, the whole systemic vasculature is considered as a single block, and as a result, the internal pressure and flow rate distribution in each individual segment of the vessel network is not computed [91]. Multi-compartment models, which build a full arterial network by connecting several mono-compartment models (see Figure 2.3), have been developed to address these shortcomings [91]. In these models, each mono-compartment model is described by a combination of its own resistance, R, compliance, C, and inductance, L, depending on the local vessel characteristics, and is considered as a building block in the

development of the whole vessel network model [91]. An example of this model connecting several RLC models will be used in this work. Currently, four typical compartmental configurations in multi-compartment models exist in the literature: L network element, inverted L network element, T element and Π element [91, 121, 138].

In practice multi-compartment models can provide a detailed and accurate description of the CVS. However, in general it is very difficult to collect all model parameter information from observations/measurements to build a patient-specific model. In principle, it is never possible to have all model parameter values at hand from measurements, since knowing more parameters requires more measurements. However, estimation is possible, if the model is reliable and if the importance rank of model parameters in the CVS is identified. In this regard, mathematical tools necessary for good modeling practices and for estimating CVS model parameters are very useful.

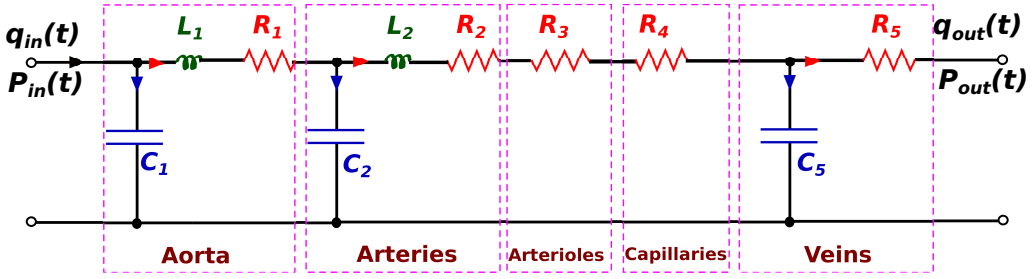

Figure 2.3: A sample multi-compartment model of the systemic circulation. Adapted from [91, 93].

2.3 Main mathematical tools used

Nowadays, it is obvious that computational cardiovascular models have become very useful for studying and analyzing the physiological and pathological processes as well as the behavior of the whole (or part) of the CVS. It is also recognized that cardiovascular models used together with mathematical tools play a crucial role in designing and constructing patient-specific models for multiple purposes, including

noninvasive parameter estimation using clinically obtained measurements.

2.3.1 Discrete ensemble Kalman filter

One of the main mathematical tools used in this work is the ensemble Kalman filter (EnKF). However, before talking about the discrete EnKF, it is necessary to first recall what the discrete Kalman filter is.

2.3.1.1 Discrete Kalman filter

Named after Rudolf E. Kálmán, one of the primary developers of its theory [139], the Kalman filter (KF), in discrete context, is an algorithm suitable for estimating in a recursive manner the internal state (i.e., unknown parameters) of a linear dynamic system, from its previous state, the commands applied, and a series of noisy/inaccurate measurements observed. This means that only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state in the KF [140]. In contrast to batch estimation techniques, no history of observations and/or estimates is required [140]. In other words, the KF, in discrete context, is a recursive estimator in which observations are used as they become available to update the present state of the model [141]. This estimator works in a two-step process [140]:

- a **prediction phase** where an *a priori* state estimate (the *forecast* state estimate) is produced using the state estimate from the previous time step and without including the observation (measurement) information from the current time step;

- and an **update phase** where the current *forecast* state estimate is refined/corrected/updated using the available observation information to produce an *a posteriori* state estimate (improved estimate of the state also known as updated state or *analysis* state).

To be more precise, let us go into a little more detail about this process based on a typical linear system involving the use of KF. Let us consider a physical model (controlled process) governed by a discrete linear stochastic-dynamic system represented in matrix-vector notation as follows:

$$\mathbf{x}_k = \mathbf{A}_k\mathbf{x}_{k-1} + \mathbf{B}_k\mathbf{u}_k + \mathbf{w}_k \quad \text{(state Equation)}, \tag{2.3a}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \quad \text{(measurement Equation)}, \tag{2.3b}$$

for the discrete time steps $k = 1, 2, \ldots$, where $\mathbf{x}_k \in \mathbb{R}^n$ represents the true state vector, $\mathbf{A}_k \in \mathbb{R}^{n\times n}$ is the true dynamics (transition) matrix, $\mathbf{B}_k \in \mathbb{R}^{n\times p}$ is the control-input matrix, $\mathbf{u}_k \in \mathbb{R}^p$ is the control vector, $\mathbf{y}_k \in \mathbb{R}^m$ is the model measurement vector related to the state, $\mathbf{x}_k$, through the measurement/observation matrix, $\mathbf{H}_k \in \mathbb{R}^{m\times n}$, m, n, and p are the dimensions of the model measurement vector, the model state vector, and the control vector, respectively. The model noise, $\mathbf{w}_k \in \mathbb{R}^n$, and the observation noise, $\mathbf{v}_k \in \mathbb{R}^m$, are assumed to be zero-mean Gaussian white noises with known covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively, at time t_k. Denoting the expectation and the transpose operators by the symbol $\mathbb{E}$ and the superscript $\top$, respectively, this assumption means in mathematical terms that

- $\mathbb{E}[\mathbf{w}_k] = 0$, $\mathbb{E}[\mathbf{v}_k] = 0$, and $\forall\ k \neq l\ \ \mathbb{E}[\mathbf{w}_k\mathbf{w}_l^\top] = 0$, $\mathbb{E}[\mathbf{v}_k\mathbf{v}_l^\top] = 0$ (white noises),
- $\mathbb{E}[\mathbf{w}_k\mathbf{w}_k^\top] = \mathbf{Q}_k$ and $\mathbb{E}[\mathbf{v}_k\mathbf{v}_k^\top] = \mathbf{R}_k$ (covariance matrices of $\mathbf{w}_k$ and $\mathbf{v}_k$),
- $\mathbb{E}[\mathbf{w}_k\mathbf{v}_k^\top] = 0$ and $\mathbb{E}[\mathbf{v}_k\mathbf{w}_k^\top] = 0$ (independent noises),
- $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$ and $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k)$ (zero-mean Gaussian noises with covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$).

The KF problem statement is: *given a prior (forecast or background) estimate* $\mathbf{x}_k^{\mathrm{f}}$ *of the system state at time* t_k*, what is the update (or analysis)* $\mathbf{x}_k^{\mathrm{a}}$ *based on the observations (or measurements)* $\mathbf{y}_k$? More precisely, at time t_k, the KF aims at finding an updated state $\mathbf{x}_k^{\mathrm{a}}$ of the *forecast* estimate $\mathbf{x}_k^{\mathrm{f}}$ minimizing $\mathbf{P}_k^{\mathrm{a}}$, the covariance of the *analysis* error

$\mathbf{e}_k^{\mathrm{a}}$ (the error between the updated state $\mathbf{x}_k^{\mathrm{a}}$ and the true state $\mathbf{x}_k$) with

$$\mathbf{e}_k^{\mathrm{a}} = \mathbf{x}_k - \mathbf{x}_k^{\mathrm{a}}, \tag{2.4a}$$

$$\mathbf{P}_k^{\mathrm{a}} = \mathbb{E}\left[(\mathbf{e}_k^{\mathrm{a}})(\mathbf{e}_k^{\mathrm{a}})^\top\right]. \tag{2.4b}$$

2.3.1.1.1 Prediction phase

As shown in Figure 2.4, in the KF procedure, the prediction phase mainly consists of two steps: the prediction $\mathbf{x}_k^{\mathrm{f}}$ of the current state (also known as the *a priori* state estimate or the *forecast* state estimate), and the estimation of the *forecast* error covariance $\mathbf{P}_k^{\mathrm{f}}$ (i.e., the covariance of the error between the *forecast* state and the true state).

2.3.1.1.1.1 Prediction $\mathbf{x}_k^{\mathrm{f}}$ of the current state

At time step k, the current state is predicted by adapting the state equation in the dynamical system (Equation (2.3a)) to produce a *forecast* model as

$$\mathbf{x}_k^{\mathrm{f}} = \mathbf{A}_k\mathbf{x}_{k-1}^{\mathrm{a}} + \mathbf{B}_k\mathbf{u}_k, \tag{2.5}$$

where the superscripts a and f denote *analysis* and *forecast*, respectively (see Figure 2.4). This *forecast* model is an alternative way of predicting the state of the system at any given time step k.

2.3.1.1.1.2 Estimation of $\mathbf{P}_k^{\mathrm{f}}$, the covariance of the *forecast* error $\mathbf{e}_k^{\mathrm{f}}$

In this step, like Equations (2.4a) and (2.4b) above, the *forecast* error, $\mathbf{e}_k^{\mathrm{f}}$, and its covariance, $\mathbf{P}_k^{\mathrm{f}}$, are defined as

$$\mathbf{e}_k^{\mathrm{f}} = \mathbf{x}_k - \mathbf{x}_k^{\mathrm{f}}, \tag{2.6a}$$

$$\mathbf{P}_k^{\mathrm{f}} = \mathbb{E}\left[(\mathbf{e}_k^{\mathrm{f}})(\mathbf{e}_k^{\mathrm{f}})^\top\right]. \tag{2.6b}$$

Thus, at time t_k, using Equations (2.6a), (2.3a), (2.5), and (2.4a), the *forecast* error can be rewritten as

$$\begin{aligned}\mathbf{e}_k^{\mathrm{f}} &= (\mathbf{A}_k\mathbf{x}_{k-1} + \mathbf{B}_k\mathbf{u}_k + \mathbf{w}_k) - (\mathbf{A}_k\mathbf{x}_{k-1}^{\mathrm{a}} + \mathbf{B}_k\mathbf{u}_k)\\ &= \mathbf{A}_k\mathbf{e}_{k-1}^{\mathrm{a}} + \mathbf{w}_k,\end{aligned} \tag{2.7}$$

where $\mathbf{e}^{\mathrm{a}}_{k-1}$ is the *analysis* error at previous time t_{k-1}.

Then, assuming that the model error and the *analysis* error are uncorrelated over time (i.e., $\forall\ k, l\ \ \mathbb{E}[(\mathbf{e}^{\mathrm{a}}_k)(\mathbf{w}_l)^\top] = \mathbb{E}[(\mathbf{w}_l)(\mathbf{e}^{\mathrm{a}}_k)^\top] = 0$) and using Equations (2.6b), (2.7), and (2.4b), the *forecast* error covariance $\mathbf{P}^{\mathrm{f}}_k$ can be obtained as

$$\begin{aligned}\mathbf{P}^{\mathrm{f}}_k &= \mathbb{E}[(\mathbf{A}_k\mathbf{e}^{\mathrm{a}}_{k-1} + \mathbf{w}_k)(\mathbf{A}_k\mathbf{e}^{\mathrm{a}}_{k-1} + \mathbf{w}_k)^\top] \\ &= \mathbf{A}_k\mathbb{E}[(\mathbf{e}^{\mathrm{a}}_{k-1})(\mathbf{e}^{\mathrm{a}}_{k-1})^\top]\mathbf{A}^\top_k + \mathbf{A}_k\mathbb{E}[(\mathbf{e}^{\mathrm{a}}_{k-1})(\mathbf{w}_k)^\top] + \mathbb{E}[(\mathbf{w}_k)(\mathbf{e}^{\mathrm{a}}_{k-1})^\top]\mathbf{A}^\top_k + \mathbb{E}[(\mathbf{w}_k)(\mathbf{w}_k)^\top] \\ &= \mathbf{A}_k\mathbf{P}^{\mathrm{a}}_{k-1}\mathbf{A}^\top_k + \mathbf{Q}_k.\end{aligned} \tag{2.8}$$

2.3.1.1.2 Update phase

This last phase in the KF procedure mainly works in three steps. As shown in Figure 2.4, using the available observation information, it starts with the Kalman gain computation step and ends with the *forecast* error covariance update (or *analysis* error covariance computation) step passing through the *forecast* state update (or *analysis* state computation) step. This is the usual order of these three steps as they appear in the KF update phase. However, in order to better explain the derivation of the different equations involved in this phase, this standard order will be disrupted in the description of these three updating steps below.

2.3.1.1.2.1 *Forecast* state update step

In this phase where the available observation information is used, an estimate of the state (the *a posteriori* state estimate $\mathbf{x}^{\mathrm{a}}_k$) at time t_k is sought by assuming it to be a linear combination of the *forecast* state, given by Equation (2.5), and the observations, given by Equation (2.3b) [142],

$$\mathbf{x}^{\mathrm{a}}_k = \tilde{\mathbf{L}}_k\mathbf{x}^{\mathrm{f}}_k + \tilde{\mathbf{K}}_k\mathbf{y}_k, \tag{2.9}$$

where $\tilde{\mathbf{L}}_k$ and $\tilde{\mathbf{K}}_k$ are two weighting unknown matrices to be determined. For an unbiased *forecast* error (i.e. $\mathbb{E}[\mathbf{e}^{\mathrm{f}}_k] = 0$), Todling and Cohn [142] demonstrated that the

analysis error is unbiased only if the weighting matrix $\tilde{\mathbf{L}}_k$ is:

$$\tilde{\mathbf{L}}_k = \mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k. \tag{2.10}$$

As a result, the formula for the *a posteriori* state estimate given by Equation (2.9) can then be rewritten as

$$\mathbf{x}_k^{\mathrm{a}} = \mathbf{x}_k^{\mathrm{f}} + \tilde{\mathbf{K}}_k \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^{\mathrm{f}} \right). \tag{2.11}$$

where the term $\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^{\mathrm{f}}$ is usually known as *innovation* or *measurement pre-fit residual* [140]. However, it is worth noting that Equation (2.11) above is for any choice of *gain* matrix $\tilde{\mathbf{K}}_k$. Therefore, it is the optimal formula for the state update only for a particular choice of *gain* matrix $\tilde{\mathbf{K}}_k$ whose the derivation will be discussed below. Once this derivation is complete, the optimal formula for the state updating can finally be given.

2.3.1.1.2.2 *Forecast* error covariance update step

In this step, the *analysis* error is first obtained by subtracting $\mathbf{x}_k$ from both sides of Equation (2.11) and using Equations (2.4a) and (2.6a). This gives

$$\mathbf{e}_k^{\mathrm{a}} = \mathbf{e}_k^{\mathrm{f}} - \tilde{\mathbf{K}}_k \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^{\mathrm{f}} \right). \tag{2.12}$$

This means that at observation locations, the *gain* matrix $\tilde{\mathbf{K}}_k$ describes the weights given to the *innovation* [142]. By substituting then Equation (2.3b) into Equation (2.12), the *analysis* error becomes

$$\mathbf{e}_k^{\mathrm{a}} = (\mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k) \mathbf{e}_k^{\mathrm{f}} - \tilde{\mathbf{K}}_k \mathbf{v}_k. \tag{2.13}$$

Finally, using Equations (2.4b) and (2.13) the *analysis* error covariance matrix can be written as[142].

$$\mathbf{P}_k^{\mathrm{a}} = (\mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k) \mathbf{P}_k^{\mathrm{f}} (\mathbf{I} - \tilde{\mathbf{K}}_k \mathbf{H}_k)^{\top} + \tilde{\mathbf{K}}_k \mathbf{R}_k \tilde{\mathbf{K}}_k^{\top} \tag{2.14}$$

where $\mathbf{R}_k = \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^{\top}]$ is the observation error covariance. Again, it is worth noting that Equation (2.14) above describes the evolution of the *analysis* error covariance for any choice of *gain* matrix $\tilde{\mathbf{K}}_k$ and is therefore not the optimal formula for updating the error

covariance. The optimal formula for the error covariance update will be given below once the particular choice of the *gain* matrix $\tilde{\mathbf{K}}_k$ has been made.

2.3.1.1.2.3 Optimal *gain* matrix computation step

The optimal *gain* matrix $\mathbf{K}_k$ is obtained from a particular choice of the *gain* matrix $\tilde{\mathbf{K}}_k$ consisting in finding the *gain* matrix $\tilde{\mathbf{K}}_k$ minimizing the *trace* of the *analysis* error covariance matrix. In fact, the diagonal elements of the *analysis* error covariance matrix are the mean squared errors. Therefore, minimizing the *trace* of $\mathbf{P}_k^{\mathrm{a}}$ amounts to minimizing the sum of the mean squared errors.

Thus, by expanding out and collecting the terms in Equation (2.14), the formula for updating the error covariance can be rewritten as

$$\mathbf{P}_k^{\mathrm{a}} = \mathbf{P}_k^{\mathrm{f}} - \mathbf{P}_k^{\mathrm{f}}\mathbf{H}_k^{\top}\tilde{\mathbf{K}}_k^{\top} - \tilde{\mathbf{K}}_k\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}} + \tilde{\mathbf{K}}_k\left(\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{H}_k^{\top} + \mathbf{R}_k\right)\tilde{\mathbf{K}}_k^{\top}. \tag{2.15}$$

Applying then the *trace* operator to the formula (2.15) and using the properties of *trace* operator to collect some terms, this becomes

$$\mathrm{tr}\left(\mathbf{P}_k^{\mathrm{a}}\right) = \mathrm{tr}\left(\mathbf{P}_k^{\mathrm{f}}\right) - 2\,\mathrm{tr}\left(\tilde{\mathbf{K}}_k\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}}\right) + \mathrm{tr}\left[\tilde{\mathbf{K}}_k\left(\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{H}_k^{\top} + \mathbf{R}_k\right)\tilde{\mathbf{K}}_k^{\top}\right]. \tag{2.16}$$

where tr denotes the *trace* operator. By differentiating $\mathrm{tr}(\mathbf{P}_k^{\mathrm{a}})$ in Equation (2.16) with respect to $\tilde{\mathbf{K}}_k$ using the gradient matrix rules and the symmetry property of the matrices involved and applying finally the first order optimality condition (i.e., equating result to zero), we get

$$\frac{\partial\ \mathrm{tr}(\mathbf{P}_k^{\mathrm{a}})}{\partial\ \tilde{\mathbf{K}}_k} = -2\left(\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}}\right)^{\top} + 2\tilde{\mathbf{K}}_k\left(\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{H}_k^{\top} + \mathbf{R}_k\right) = 0. \tag{2.17}$$

From Equation (2.17) and due to the symmetry property of $\mathbf{P}_k^{\mathrm{f}}$, the minimum value of $\mathrm{tr}\left(\mathbf{P}_k^{\mathrm{a}}\right)$ is obtained when the *gain* matrix is [143]

$$\tilde{\mathbf{K}}_k = \mathbf{K}_k = \mathbf{P}_k^{\mathrm{f}}\mathbf{H}_k^{\top}\left(\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{H}_k^{\top} + \mathbf{R}_k\right)^{-1}. \tag{2.18}$$

This matrix $\mathbf{K}_k$ is the optimal weighting matrix referred to as the **Kalman gain**. The term $\mathbf{H}_k\mathbf{P}_k^{\mathrm{f}}\mathbf{H}_k^{\top} + \mathbf{R}_k$ included in $\mathbf{K}_k$, commonly referred to as *innovation covariance*

or *measurement pre-fit residual covariance* and usually denoted by $\mathbf{S}_k$ [140] is the measurement prediction covariance associated to the *innovation* term $\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k^{\mathrm{f}}$ defined above in the state update step (Section 2.3.1.1.2.1). Now the optimal *gain* matrix is calculated, the optimal formulas for updating the state and the error covariance can be provided.

2.3.1.1.2.4 Optimal formulas for updating the error covariance and state

Substituting the Kalman gain into Equation (2.14) yields this simplified and optimal formula for *analysis* error covariance matrix

$$\mathbf{P}_k^{\mathrm{a}} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\,\mathbf{P}_k^{\mathrm{f}}. \tag{2.19}$$

Finally, by replacing in Equation (2.11) the weighting matrix $\tilde{\mathbf{K}}_k$ by the optimal *gain* matrix $\mathbf{K}_k$, the optimal *analysis* state becomes

$$\mathbf{x}_k^{\mathrm{a}} = \mathbf{x}_k^{\mathrm{f}} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k^{\mathrm{f}}\right). \tag{2.20}$$

These different stages of the KF algorithmic loop discussed above are summarized in the diagram of Figure 2.4. However, it is worth noting that like the implementation of most algorithms, the implementation of the KF requires an initialization step which plays an important role to obtain desired performance. Hence, in the classical KF, an initial guess of state estimate, $\mathbf{x}_0^{\mathrm{a}}$, and an initial guess of the error covariance matrix, $\mathbf{P}_0^{\mathrm{a}}$, must be first provided as initial values by the user before proceeding with anything else. After this initialization phase, the different stages of the KF presented above can be easily implemented for each time step, $k = 1, 2, 3, \ldots$. These traditional KF equations serve as the mathematical foundation for the common variants of the KF, including the extended Kalman filter, the unscented Kalman filter, and the ensemble Kalman filter.

2.3.1.2 Ensemble Kalman filter algorithm

As described in Section 2.3.1.1, the KF addresses the general problem of trying to estimate the state of a discrete, linear, and controlled stochastic-dynamical system,

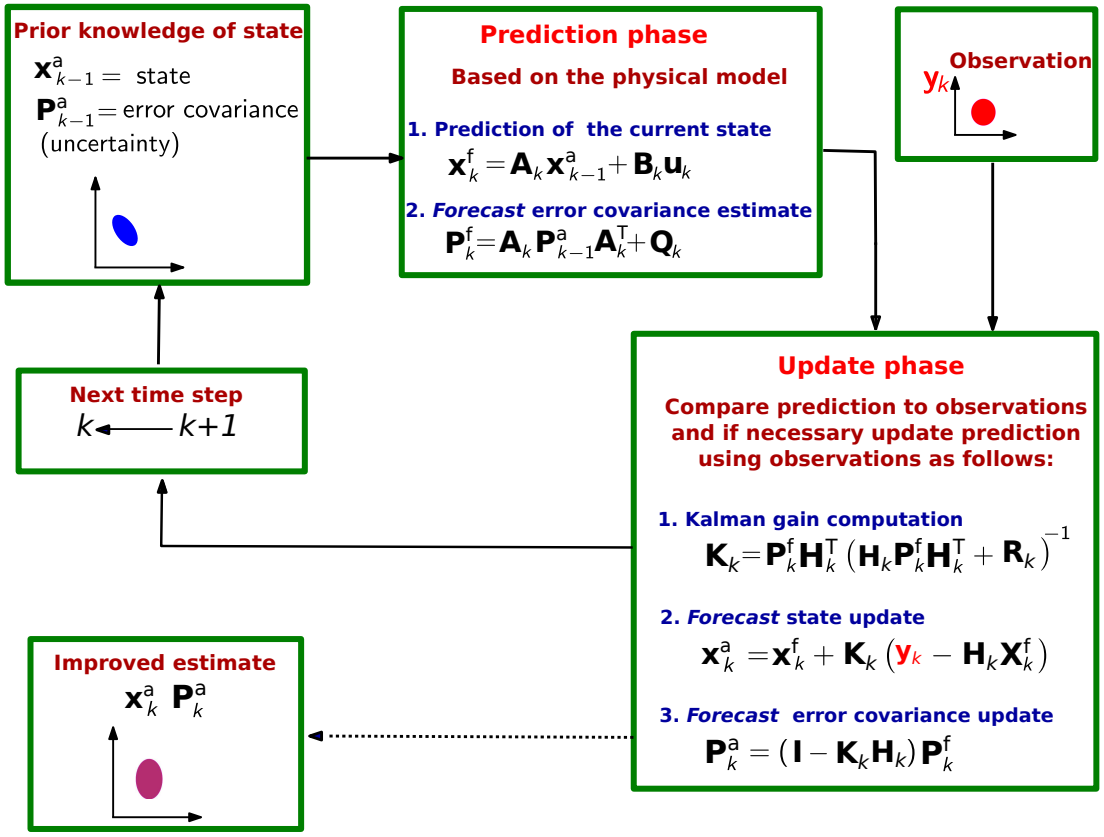

Figure 2.4: A schematic representation of the Kalman filter algorithmic loop.

process or model. But what happens if, on the one hand, the process to be estimated and (or) the observation relationship to the process is nonlinear, and if, on the other hand, the state and observation dimensions are huge? In such situations, the combination of high dimensionality and nonlinearity makes this a very challenging problem. In this regard, several variants of the standard KF have been proposed in the literature to meet this challenge [141, 144, 145, 146, 147, 148, 149, 150]. One of them is the ensemble Kalman filter (EnKF). First introduced by Geir Evensen [141] and later clarified by Burgers et al. [147], the EnKF is a sequential Monte Carlo approximation of the KF, which avoids evolving the covariance matrix of the state vector distribution [151]. Instead, it uses an ensemble of model states integrated forward in time with a nonlinear model including replicates of system noise for estimating the true states of the model [152]. Indeed, as outlined by Katzfuss et al., in

contrast to the standard KF, which works with the entire distribution of the state explicitly, the EnKF only considers a sample (an ensemble of vectors) approximating the overall state distribution [149]. This sample (ensemble) is then propagated forward through time and updated when new observations are available. In general, this ensemble representation of the state space is viewed as a form of dimension reduction, in that only a small ensemble is propagated instead of the joint distribution including the full covariance matrix [149].

Furthermore, the EnKF is also recognized today as a sub-optimal estimator for problems involving high-order nonlinear models. In this regard, as so aptly stated by Katzfuss et al. [149], quoting Tukey [150], the EnKF embodies the principle that an approximate solution to the right problem is worth more than a precise solution to the wrong problem. As such, for many realistic, highly complex systems, the EnKF is essentially the only way to do (approximate) inference, while alternative exact inference techniques can only be applied to highly simplified versions of the problem [149]. Also, the advantage of dimensional reduction in the EnKF leads to computational facilities and feas even for very high-dimensional systems. Below, the different steps employed in the formulation of EnKF as presented in Lal et al. [153] are discussed.

2.3.1.2.1 Derivation of the EnKF

Like the traditional KF, the EnKF consists of a prediction phase and an update phase at each time step. As a result, the formulation of EnKF is very similar to the KF one except for the following.

2.3.1.2.1.1 Nonlinear system

First, for the EnKF derivation, the discrete-time linear dynamical system — Equations (2.3a) and (2.3b) — used for the KF derivation is typically replaced with the following discrete nonlinear system :

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k, \quad \mathbf{y}_k = h(\mathbf{x}_k) + \mathbf{v}_k. \tag{2.21}$$

where the nonlinear functions f and h are the state transition and observation models, respectively. Other variables have the same definitions as in Section 2.3.1.1. Specifically, $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{y}_k \in \mathbb{R}^{nobs}$, and $\mathbf{u}_k \in \mathbb{R}^m$ are, at time t_k, the model state, the observed measurement, and the control parameter, respectively. Therefore, n, $nobs$, and m represent the dimension of the model state vector, the number of observations, and the number of model parameters, respectively. $\mathbf{w}_k \in \mathbb{R}^n$ and $\mathbf{v}_k \in \mathbb{R}^{nobs}$ are assumed uncorrelated Gaussian model errors with $\mathbf{w}_k \sim \mathcal{N}(0, \mathbf{Q_k})$ and $\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R_k})$ where $\mathbf{Q_k}$ and $\mathbf{R_k}$ are the covariance matrices.

2.3.1.2.1.2 Initial *forecast* ensemble of states

Then, in the prediction phase, the initial *forecast* ensemble of states (prior ensembles) $\mathbf{X}_k^{\mathrm{f}} = (\mathbf{x}_k^{\mathrm{f}_1}, \ldots, \mathbf{x}_k^{\mathrm{f}_{q_{\mathrm{ens}}}}) \in \mathbb{R}^{n \times q_{\mathrm{ens}}}$ for $i = 1, \ldots, q_{\mathrm{ens}}$ is assumed to be available at time step k. f_i stands for the initial i^{th} *forecast* member of an ensemble size q_{ens}. The mean of the ensemble of *forecast* state is $\overline{\mathbf{x}}_k^{\mathrm{f}} \in \mathbb{R}^n$ and is given by

$$\overline{\mathbf{x}}_k^{\mathrm{f}} = \frac{1}{q_{\mathrm{ens}}} \sum_{i=1}^{q_{\mathrm{ens}}} \mathbf{x}_k^{\mathrm{f}_i}. \tag{2.22}$$

2.3.1.2.1.3 Error covariance matrices computation

The error covariance matrices for the *forecast* and the *analyzed* estimate, $\mathbf{P}_k^{\mathrm{f}}$ and $\mathbf{P}_k^{\mathrm{a}}$, are not defined in terms of the true state as Equations (2.6b) and (2.4b) in the standard KF. In fact, in the EnKF, the true state is not known. Therefore, according to the interpretation that the ensemble mean is the best estimate and the spreading of the ensemble around the mean is a natural definition of the error in the ensemble mean [154], the error covariance matrices are defined around the ensemble mean as follows:

$$\mathbf{P}_k^{\mathrm{f}} = \frac{1}{q_{\mathrm{ens}} - 1} \sum_{i=1}^{q_{\mathrm{ens}}} \left(\mathbf{x}_k^{\mathrm{f}_i} - \overline{\mathbf{x}}_k^{\mathrm{f}} \right) \left(\mathbf{x}_k^{\mathrm{f}_i} - \overline{\mathbf{x}}_k^{\mathrm{f}} \right)^\top, \tag{2.23}$$

$$\mathbf{P}_k^{\mathrm{a}} = \frac{1}{q_{\mathrm{ens}} - 1} \sum_{i=1}^{q_{\mathrm{ens}}} \left(\mathbf{x}_k^{\mathrm{a}_i} - \overline{\mathbf{x}}_k^{\mathrm{a}} \right) \left(\mathbf{x}_k^{\mathrm{a}_i} - \overline{\mathbf{x}}_k^{\mathrm{a}} \right)^\top. \tag{2.24}$$

2.3.1.2.1.4 Update of the ensemble of states

After the Kalman gain ($\mathbf{K}_k$) computation, all operations on the ensemble members are

independent in the EnKF update phase and the ensemble members are updated using:

$$\mathbf{x}_k^{\mathrm{a}_i} = \mathbf{x}_k^{\mathrm{f}_i} + \mathbf{K}_k \left[\mathbf{y}_k^i - h\left(\mathbf{x}_k^{\mathrm{f}_i}\right) \right], \quad i = 1, \ldots, q_{\mathrm{ens}}, \tag{2.25}$$

where a_i represents the i^{th} updated or analyzed member of the ensemble. One of the EnKF specificities is that the observations are treated as random variables by generating an ensemble of observations using small perturbations to avoid having an updated ensemble with a low variance [147]. These perturbations are generated using the same distribution as the measurement error. This special treatment given to observations permits to preserve from one EnKF iteration to the next the correct *forecast* error covariance [126]. In concrete terms, small perturbations $\mathbf{e}_k^i$, $i = 1, \ldots, q_{\mathrm{ens}}$ are added to the original observation set $\mathbf{y}_k$ in order to generate an ensemble of q_{ens} perturbed observations $\mathbf{y}_k^i$ $i = 1, \ldots, q_{\mathrm{ens}}$ as follows:

$$\mathbf{y}_k^i = \mathbf{y}_k + \mathbf{e}_k^i, \quad i = 1, \ldots, q_{\mathrm{ens}} \tag{2.26}$$

with $\mathbf{e}_k^i \in \mathbb{R}^{nobs}$ and $\mathbf{e}_k^i \sim \mathcal{N}(0, \mathbf{R_k})$. The measurement error covariance matrix, $\mathbf{R_k}$, is diagonal following the assumption of independent observations [155] and is defined as

$$\mathbf{R_k} = \mathrm{diag}\left[\frac{1}{q_{\mathrm{ens}} - 1} \mathbf{E}\mathbf{E}^\top \right], \quad \mathbf{E} = \left[\mathbf{e}_k^1, \ldots, \mathbf{e}_k^{q_{\mathrm{ens}}} \right]. \tag{2.27}$$

Thus, at each time step of the EnKF procedure, the observations are treated as random variables.

2.3.1.2.1.5 Kalman gain computation

If the measurement function h is linear and the noise is additive, namely,

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v_k}, \tag{2.28}$$

the Kalman gain is given by [156]

$$\mathbf{K_k} = \mathbf{P}_k^{\mathrm{f}} \mathbf{H}^\top \left(\mathbf{H}\mathbf{P}_k^{\mathrm{f}} \mathbf{H}^\top + \mathbf{R}_k \right)^{-1}, \tag{2.29}$$

where $\mathbf{H} \in \mathbb{R}^{nobs \times n}$, is a linear or linearized observation operator. To avoid having to resort to the linearization of a nonlinear measurement function which could sometimes

be difficult to linearize, Houtekamer and Mitchell [157] rewrote the two terms $\mathbf{P}_k^{\mathrm{f}}\mathbf{H}^\top$ and $\mathbf{H}\mathbf{P}_k^{\mathrm{f}}\mathbf{H}^\top$ which appear in the Kalman gain (Equation (2.29)) as

$$\mathbf{P}_k^{\mathrm{f}}\mathbf{H}^\top \equiv \frac{1}{q_{\mathrm{ens}}-1}\sum_{i=1}^{q_{\mathrm{ens}}}\left[\mathbf{x}_k^{\mathrm{f}_i}-\overline{\mathbf{x}}_k^{\mathrm{f}}\right]\left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)-\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]^\top, \tag{2.30}$$

$$\mathbf{H}\mathbf{P}_k^{\mathrm{f}}\mathbf{H}^\top \equiv \frac{1}{q_{\mathrm{ens}}-1}\sum_{i=1}^{q_{\mathrm{ens}}}\left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)-\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]\left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)-\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]^\top, \tag{2.31}$$

where $\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)} = \frac{1}{q_{\mathrm{ens}}}\sum_{i=1}^{q_{\mathrm{ens}}} h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)$. Ambadan and Tang [158] argued that Equations (2.30) and (2.31) approximately hold only if the following conditions are fulfilled:

$$\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)} = h\left(\overline{\mathbf{x}}_k^{\mathrm{f}}\right), \tag{2.32}$$

$$\mathrm{norm}\left(\mathbf{x}_k^{\mathrm{f}_i}-\overline{\mathbf{x}}_k^{\mathrm{f}}\right) \text{ is small for } i=1,2,\ldots q_{\mathrm{ens}}. \tag{2.33}$$

In fact, as argued by Ambadan and Tang [158], under the conditions of Equations (2.32) and (2.33), Equations (2.30) and (2.31) linearize the nonlinear function h to $\mathbf{H}$. Therefore, when the Gaussian nonlinear model is associated with a nonlinear measurement function h, the Kalman gain is defined as [159]:

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{x}\mathbf{y}_k}^{\mathrm{f}}\left(\mathbf{P}_{\mathbf{y}\mathbf{y}_k}^{\mathrm{f}}\right)^{-1}, \tag{2.34}$$

where $\mathbf{P}_{\mathbf{x}\mathbf{y}_k}^{\mathrm{f}}$ is the cross-covariance between the state and observation errors and $\mathbf{P}_{\mathbf{y}\mathbf{y}_k}^{\mathrm{f}}$, the error covariance of the difference between the observation and its prediction. They can be obtained as follows:

$$\mathbf{P}_{\mathbf{x}\mathbf{y}_k}^{\mathrm{f}} = \frac{1}{q_{\mathrm{ens}}-1}\sum_{i=1}^{q_{\mathrm{ens}}}\left[\mathbf{x}_k^{\mathrm{f}_i}-\overline{\mathbf{x}}_k^{\mathrm{f}}\right]\left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)-\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]^\top, \tag{2.35}$$

$$\mathbf{P}_{\mathbf{y}\mathbf{y}_k}^{\mathrm{f}} = \frac{1}{q_{\mathrm{ens}}-1}\sum_{i=1}^{q_{\mathrm{ens}}}\left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)-\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]\left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)-\overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]^\top + \mathbf{R}_k. \tag{2.36}$$

2.3.1.2.1.6 Prediction step

The last step is the prediction step which involves predicting an ensemble of q_{ens} forecast states for the next time step $k+1$ as,

$$\mathbf{x}_{k+1}^{\mathrm{f}_i} = f(\mathbf{x}_k^{\mathrm{a}_i}, \mathbf{u}_{k+1}) + \mathbf{w}_{k+1}^i, \;\; i=1,2,\ldots q_{\mathrm{ens}}. \tag{2.37}$$

2.3.1.2.2 Summary of the ensemble Kalman filter algorithm

This Section summarizes the *forecast* and the *analysis* steps of EnKF presented in Section 2.3.1.2.1. But before that, a schematic description of the EnKF algorithm is shown in Figure 2.5. To start the EnKF, it is necessary first to generate an ensemble of q_{ens} *forecast* estimates of state associated with their random errors. Thus, at time step $k-1$, the ensemble of *analysis* state estimates, $\mathbf{x}_{k-1}^{\mathrm{a}_i}$ for $i = 1, \ldots, q_{\text{ens}}$, is assumed to be available. Therefore, at time step k, an ensemble of *forecast* state estimates, $(\mathbf{x}_k^{\mathrm{f}_1}, \ldots, \mathbf{x}_k^{\mathrm{f}_{q_{\text{ens}}}})$ is generated. The corresponding measurements for this ensemble of states are denoted as $(\mathbf{y}_k^1, \ldots, \mathbf{y}_k^{q_{\text{ens}}}) \in \mathbb{R}^{q_{\text{ens}} \times nobs}$ where *nobs* is the number of observations. $\mathbf{Q}_k$ and $\mathbf{R}_k$ correspond to the model and observation error covariance matrices, respectively, at time step k. Finally, given an *analysis* ensemble $(\mathbf{x}_{k-1}^{\mathrm{a}_1}, \ldots, \mathbf{x}_{k-1}^{\mathrm{a}_{q_{\text{ens}}}})$ at time step k, the equations for the EnKF can be written as:

$$
\begin{aligned}
\mathbf{x}_k^{\mathrm{f}_i} &= f(\mathbf{x}_{k-1}^{\mathrm{a}_i}, \mathbf{u}_k) + \mathbf{w}_k^i, \;\; i = 1, \ldots, q_{\text{ens}}, \\
\mathbf{w}_k^i &\sim \mathcal{N}(0, \mathbf{Q}_k), \\
\mathbf{P}_{\mathbf{x}\mathbf{y}_k}^{\mathrm{f}} &= \frac{1}{q_{\text{ens}} - 1} \sum_{i=1}^{q_{\text{ens}}} \left[\mathbf{x}_k^{\mathrm{f}_i} - \overline{\mathbf{x}}_k^{\mathrm{f}}\right] \left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right) - \overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]^{\top}, \\
\mathbf{P}_{\mathbf{y}\mathbf{y}_k}^{\mathrm{f}} &= \frac{1}{q_{\text{ens}} - 1} \sum_{i=1}^{q_{\text{ens}}} \left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right) - \overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right] \left[h\left(\mathbf{x}_k^{\mathrm{f}_i}\right) - \overline{h\left(\mathbf{x}_k^{\mathrm{f}}\right)}\right]^{\top} + \mathbf{R}_k, \\
\mathbf{K}_k &= \mathbf{P}_{\mathbf{x}\mathbf{y}_k}^{\mathrm{f}} \left(\mathbf{P}_{\mathbf{y}\mathbf{y}_k}^{\mathrm{f}}\right)^{-1}, \\
\mathbf{y}_k^i &= \mathbf{y}_k + \mathbf{e}_k^i, \;\; i = 1, \ldots, q_{\text{ens}}, \\
\mathbf{x}_k^{\mathrm{a}_i} &= \mathbf{x}_k^{\mathrm{f}_i} + \mathbf{K}_k \left[\mathbf{y}_k^i - h\left(\mathbf{x}_k^{\mathrm{f}_i}\right)\right], \;\; i = 1, \ldots, q_{\text{ens}}.
\end{aligned}
$$

where, the superscripts 'a' and 'f' denote *analysis* and *forecast*, respectively. This algorithm provides an ensemble of analyses at time step k, which can be cycled in time.

In Lal et al. [12, 153], the usefulness of the EnKF-based technique for estimating hemodynamic parameters using both synthetic and *in vitro* data is demonstrated. Specifically, in [12], the EnKF algorithm was integrated in an inversion procedure for estimating hemodynamic model parameters, which then allowed to estimate patient-specific blood pressure in cerebral arteries (see Figure 2.6). This algorithm is

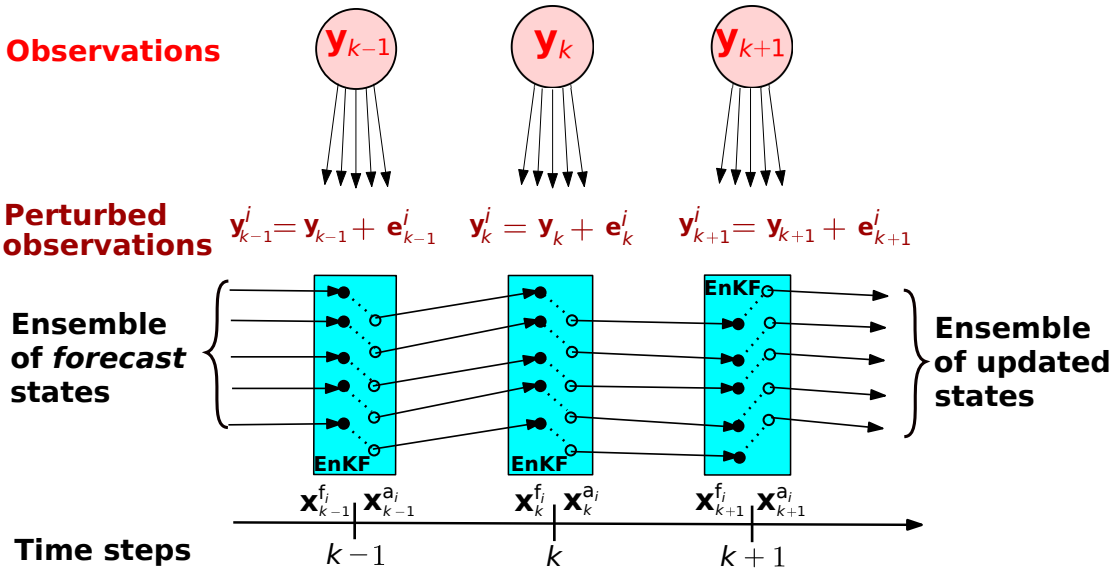

Figure 2.5: A schematic description of the ensemble Kalman filter. Adapted from [153].

reused in this work with the aim of reproducing all this inversion procedure with slight modifications to assess the inversion accuracy and robustness. In fact, from clinical data acquisition to the computer simulations of the computational cardiovascular models using this algorithm, via the derivation of physical laws of hemodynamics, there are several sources of uncertainty which can make the model unreliable and less predictive. Uncertainty must be reduced to achieve a reliable and predictive cardiovascular model for investigating CVDs. In this regard, uncertainty and sensitivity analysis practices are necessary.

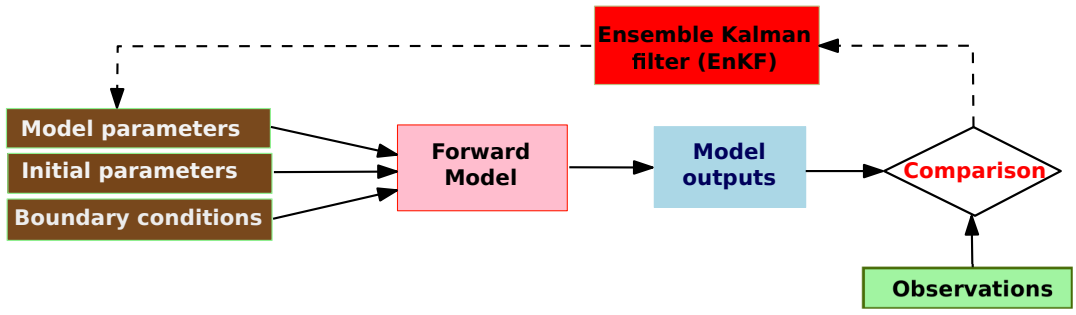

Figure 2.6: Parameter estimation flow chart using the ensemble Kalman filter. Adapted from Lal [126].

2.3.2 Uncertainty and sensitivity analysis

Uncertainty analysis (UA) and sensitivity analysis (SA) are two other important mathematical tools used in this thesis. These are two closely related mathematical practices, which should be ideally run in tandem since UA is usually a necessary precursor of SA [160]. Indeed, UA quantifies the uncertainty in the model output, while SA identifies the model input variables potentially responsible for this uncertainty. The use of these two mathematical tools in this thesis is motivated, on the one hand, by the obvious existence of various sources of uncertainty in any modeling and, on the other hand, by the effectiveness of UA and SA in the assessment, validation, and verification of simulation models and computer codes in a variety of fields.

2.3.2.1 Sources of uncertainty

Like any model, the cardiovascular models are subject to sources of uncertainty. One way to categorize these sources of uncertainty is to consider two groups of uncertainty: aleatoric uncertainty and epistemic uncertainty [161, 162]. Also known as statistical uncertainty [163], aleatoric uncertainty is uncertainty due to the occurrence of random events. It is representative of unknowns that differ each time the same experiment is run [163]. As for epistemic uncertainty (also known as systematic uncertainty), it is uncertainty due to things one could in principle know but do not in practice [163]. This may be due to [164]:

- experimental uncertainty : errors of measurement,
- interpolation uncertainty: missing information or lack of available data collected from computer model simulations and/or experimental measurements,
- numerical uncertainty: numerical errors and numerical approximations per implementation of a computer model,

- structural uncertainty: assumptions made in a model, lack of knowledge of the underlying physics and poor or partial understanding of the driving forces and mechanisms,
- parametric variability: spatio-temporal variabilities of model input parameters,
- and parameter uncertainty: model input parameters whose exact values are unknown to experimentalists and cannot be controlled in physical experiments.

Taking into account these different sources of uncertainty by performing UA can provide more information about the confidence of model results.

2.3.2.2 Uncertainty analysis

UA investigates the uncertainty of variables used in a model with the aim of improving the model results through the quantification of uncertainties in some or all model variables. In other words, UA mainly focuses on uncertainty quantification (UQ) and propagation, and often deals with assessing the uncertainty in measurement or estimation of model variables.

2.3.2.2.1 Uncertainty quantification

UQ is a technique of quantitative determination and reduction of uncertainties in both computational and real world applications. It attempts to determine the likelihood of certain outcomes if some aspects of the model or system are not exactly known [163]. There are two major types of problems in uncertainty quantification [163]: forward uncertainty propagation (where the different sources of uncertainty are propagated through the model to predict the overall uncertainty in the system response) and backward uncertainty quantification (where the model parameters are calibrated simultaneously using test data).

2.3.2.2.2 Forward uncertainty propagation

Uncertainty propagation is the quantification of uncertainty in the model outputs

(blood pressure and blood flow in a CVS model, for example) due to the effect of model input parameter uncertainties. Uncertainty in model parameters is propagated through approaches including Monte Carlo simulations, polynomial chaos expansion, perturbation method, Taylor series, etc [163]. The ultimate goal of uncertainty propagation is to evaluate the first two moments (mean and variance) and the reliability of the model outputs, to represent the model output variability in a compact way.

2.3.2.2.3 Backward uncertainty quantification: calibration, verification, validation

Backward UQ estimates the discrepancy between the observations and the mathematical model and performs parameter calibration. The latter is the process of adjusting model parameters to obtain good agreement between model predictions and experimental observations (bias correction) [165] and also of estimating the values of unknown parameters in the model if there are any [163]. Calibration is to differentiated from verification which is assessment of the model "exact" implementation and from validation which is assessment of the model response degree of agreement with the available physical observation [165].

2.3.2.3 Sensitivity analysis

Lumped-parameter models of the CVS include a large number of parameters. For reliable, predictive, and confident patient-specific cardiovascular models, the importance rank of model input parameters should be estimated with the aim of reducing the uncertainties in the model outputs by fixing, for example, less important parameters to their nominal values. This kind of process is part of SA, which is in Saltelli's words, the study of how the uncertainty in the output of a mathematical model or system (numerical or otherwise) can apportioned to different sources of uncertainty in its inputs [160, 166, 167]. SA can be subdivided into two main types: local sensitivity analysis (LSA) and global sensitivity analysis (GSA).

2.3.2.3.1 Local sensitivity analysis

LSA is one of the simplest and most common SA approaches. Easy to implement and computationally less expensive, it consists in [168]:

- Perturbing one model input parameter value around its nominal value, while keeping all other parameters fixed at their nominal values to see what effect this produces on the output [169, 170, 171, 172, 173].
- Returning the perturbed input variable to its nominal value and then repeating the procedure for all parameters one by one to capture the effect of each individual input parameter perturbation on the output.

However, despite its multiple advantages, LSA has some drawbacks. Indeed, it does not explore the effect of entire parameter spaces on output variables as well as the interactions between the parameters [174, 175]. These shortcomings can be overcome with GSA.

2.3.2.3.2 Global sensitivity analysis

Unlike LSA, GSA allows to quantify the interaction effects among the parameters and also explores the impact of entire feasible parameter spaces on output variables [160, 176, 177, 178, 179]. The only drawback of GSA is its computational cost. In this work, GSA method based on the Sobol' indices will be used. This will be applied to the patient-specific arterial network to decompose the model output variance into parts attributable to each of the model input variables.

2.3.3 Machine learning

The last mathematical tool used in this work is ML. In today's world, ML is a mathematical practice, which has gained much popularity, and whose the algorithms are employed in most of scientific field and research areas [180]. Much of the improvement on ML in research reviews is attributable to the exponential increases in computing power and bigdata processing technologies [181, 182, 183, 184]. Currently,

the widespread types of ML in the literature are supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, transduction, transfer learning, and active learning tasks [185, 186, 187, 188, 189, 190]. Each type comes with its strengths and weaknesses, but all have benefited from increased attention to procedural rigor [181]. The details of these different types of ML are beyond the scope of this thesis. However, supervised learning and transfer learning will be discussed below because of their use in this work.

2.3.3.1 Supervised learning

Supervised learning is the task of learning a function that maps input data to target labels [191]. In other words, it is a ML approach where the investigator uses a database of observations with labelled outcomes or classes. These data are generally used to develop a model to predict or classify future events, or to find which variables are most relevant to the outcomes. When labelled outcomes are continuous real number values, the supervised learning task(s) are known as regression problems, and when the labelled outcomes variables are categorical variables, the tasks are known as classification problems [191].

Common supervised learning algorithms include ordinary least squares regression [192], logistic regression [193], least absolute shrinkage and selection operator (LASSO) regression [194, 195], ridge regression [196, 197, 198], elastic net regression [199, 200, 201, 202], linear discriminant analysis [203, 204, 205], Naive Bayes classifiers [206, 207, 208, 209], support vector machines [210, 211, 212, 213, 214], Bayesian networks [215, 216, 217], a variety of decision trees [218, 219, 220] especially Random Forests [221, 222] and AdaBoost or gradient boosting classifiers [223, 224], artificial neural networks [225, 226, 227, 228, 229, 230], and ensemble methods [231, 232, 233, 234, 235]. Some of examples of supervised ML tasks include regression, classification, predictive modeling and survival analysis [182, 191]. A commonly used example is training a model to differentiate between apples, oranges

and lemons. The 'label' of each type of fruit is initially supplied to the algorithm along with the features such as colour, size, weight and shape and the algorithm learns the mix of features that differentiate the fruits. Then, when a new, 'unlabelled' fruit is presented, the model should be able to predict which type of fruit it is.

2.3.3.2 Transfer learning

Another type of ML used in this thesis, which has witnessed a great increase in popularity and is fast becoming a key instrument in many ML applications in recent years [236], is transfer learning (TL). This increase in popularity and this growing interest in TL is partly due to its ability to build accurate models in a timesaving [237]. For example, in areas such as text classification, image classification, text categorization, and computer vision, TL has been producing remarkable results [237, 238, 239, 240, 241]. However, what is TL really all about? To answer this question, let us start with a simple parents-children analogy. In a family, parents have many years of life experience that they want to pass on to their children. With all this accumulated experience, the education the children receive is a summary/digest of the parents' experience. So it can be seen as a "transfer" of knowledge from parents to the child. Another excellent analogy for TL can be this fundamental human capability to transfer knowledge from previous experiences to new circumstances [242]. In fact, TL approach works exactley in the same way. Like the tranfer process in the two previous analogies, TL addresses the problem of how to leverage previously acquired knowledge from source domains or tasks, in order to improve the efficiency and accuracy of learning in a new target domain or task that is similar [243] or different but in some way related to the original one [244, 245, 246, 247, 248]. With such a definition, it is clear that TL refers to the ability of a model or a system to recognize and interpolate, extrapolate, extract, or apply knowledge and skills learned in previous tasks to new tasks [243]. This means that, in TL, the new learning process does not start from scratch but from pre-trained model patterns learned when solving a similar or

different problem.

For example, the abilities acquired while learning to identify types of cats could apply when one learns to identify types of dogs, and knowledge gained while learning to recognize apples could apply when recognizing lemons. Another strength of TL, according to Mihalkova, lies in the fact that it is considered as one of the most effective techniques for enabling learning in situations when an adequate amount of training data for the task of interest is not available [236], as is the case in this thesis. For more mathematical definition and more details about TL, the interested reader is referred to the excellent survey of Pan and Yang [243] on TL.

2.3.3.3 Convolutional neural networks

Several pre-trained models used in TL are based on convolutional neural networks (CNNs) [249]. These are ML algorithms inspired from biological processes, such that the organization of their artificial neurons resembles that of the visual cortex of animals [250, 251]. But to better understand CNNs, one must first understand what are neural networks.

2.3.3.3.1 Neural networks

Neural networks (NNs) or artificial neural networks (ANNs) form the backbone of convolutional neural networks. They are algorithms aiming at finding an approximation of an unknown function. Indeed, formed by interconnected neurons like the human brain, they attempt to use multiple layers of calculations to imitate the concept of how the human brain interprets and draws conclusions from information. Therefore, the knowledge of human neuron function is a prerequisite for understanding ANN function.

As shown in Figure 2.7, a human neuron consists of four main elements. These are dendrites, nucleus (the heart of the neuron), soma and axon [252]. When a human being receives information, the latter behaves like electrical signals, which are received by the dendrites. Then, the electrical signals are processed by soma. This process output is

thereafter carried by the axon to the axon terminals where the output is either sent to next neurons for further processing or it is the final output. The interconnection of neurons is called neural network (NN) where electrical impulses travel around the brain. An ANN is based on the same principle. As shown in Figure 2.8, it basically works on

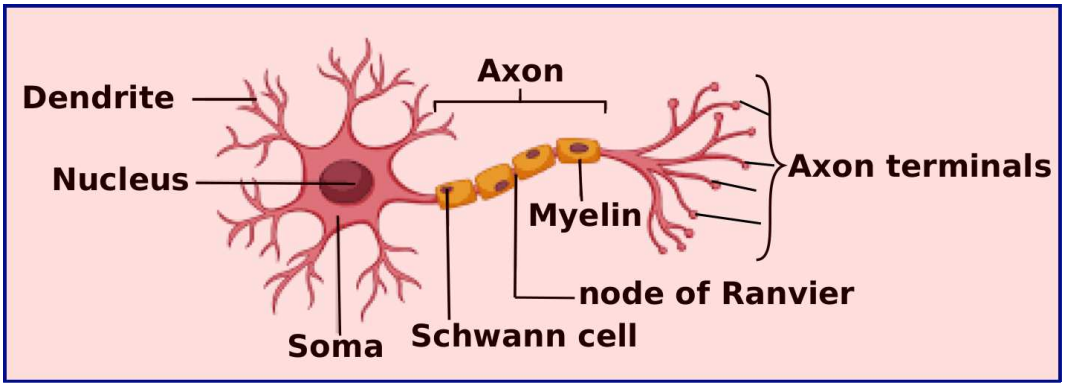

Figure 2.7: A human neuron anatomy. Adapted from [253]

three layers. The input layer takes the inputs (much like dendrites reveiving electrical signals in a human neuron). Then, the hidden layer processes the inputs and generates an output (like soma and axon). Finally, the output layer (like axon terminals) sends the calculated output either to next neurons or it is the final calculated output [252, 254]. In ML vocabulary, this process is known as **forward propagation**.

In fact, during forward propagation, each input is associated with a weight assigned to it in the hidden layer. These weights are randomly initialized and are updated during the model training process. The higher weight is assigned to the input considered more important as compared to the ones which are considered less important and a weight of zero is assigned to inputs whose the particular features are insignificant. By continually adjusting these relative weights, the network can progressively determine the best combination of weights that produce the optimal predictive accuracy. In mathematical terms, this can be written as follows:

- Let us assume the inputs to be $X_1, \cdots, X_{\mathrm{n}}$, and the weights associated to be $W_{\mathrm{k1}}, \cdots, W_{\mathrm{kn}}$, as shown in Figure 2.9.

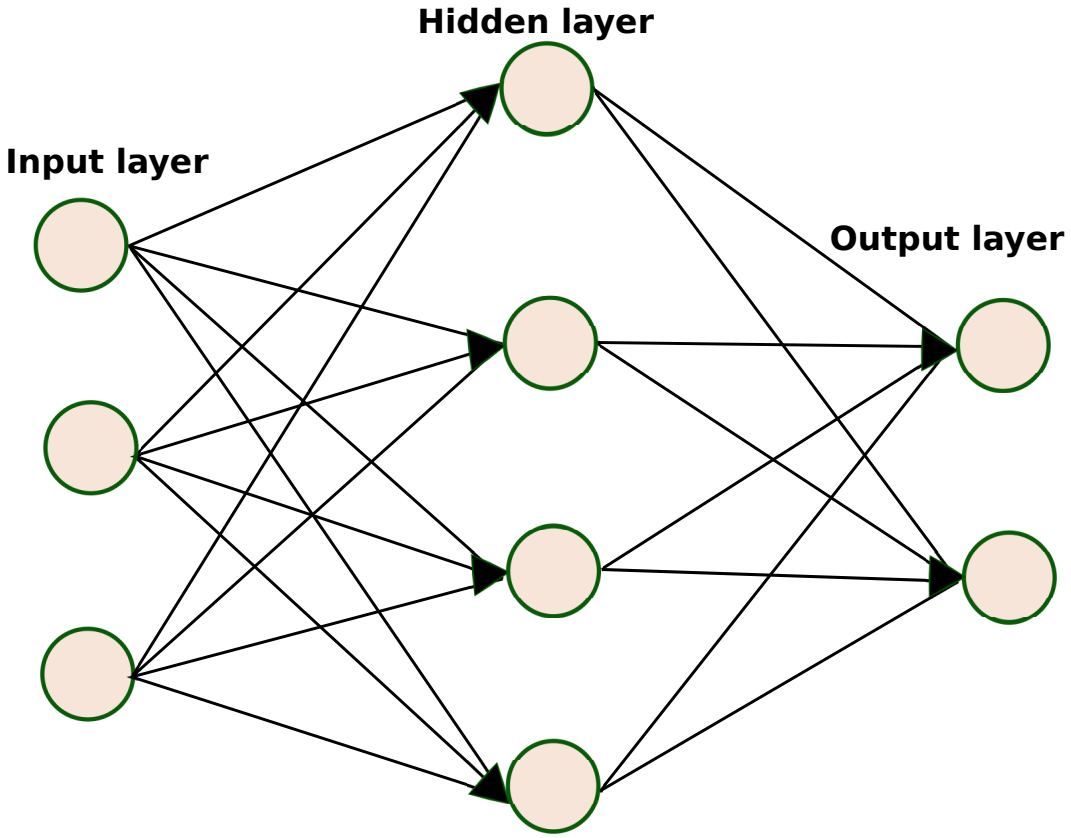

Figure 2.8: Structure of an artificial neural network. Adapted from [254]

- Passing through the nodes, the inputs are first multiplied by their corresponding weights and are then added together along with the bias, denoted by b_{k}. Let this linear component be called as u. Then,

$$u = \sum_{i=1}^{n} X_{\mathrm{i}} \times W_{\mathrm{ki}} + b_{\mathrm{k}}. \tag{2.38}$$

- Applying the activation function (a nonlinear function), to u, the final output from the neuron k is obtained as $Y_{\mathrm{k}} = f(u)$.

Sigmoid, hyperbolic tangent function, Rectified Linear Units [255], and softmax activation functions [238, 256, 257, 258, 259, 260] are the most commonly applied activation functions [261].

However, forward propagation alone does not allow the constructed network to predict the output closely to the actual value. In general, the accuracy of the network

is measured using an objective function (also referred to as a cost/loss/performance function), which tries to penalize the network when it makes errors. In fact, once the output is received for a single forward propagation iteration, the network error can be calculated. For instance, the cost function can be defined as being the mean squared error, and in this case, it can be written as :

$$C = \frac{1}{m}\sum_{i=1}^{m}(Y_{\text{obs},i} - Y_i)^2, \tag{2.39}$$

where m, Y_i, and $Y_{\text{obs},i}$ are the number of training inputs, the predicted values, and the actual values of that particular example, respectively.

The learning process objective is to adjutst the free parameters of the network (i.e., biases and weights) in order to increase the prediction accuracy (to reach the desired network output) and to reduce the error by minimizing as much as possible this cost function. In this regard, as outlined by Gupta [262], after each forward propagation iteration, the error is fed back to the network along with the gradient of the cost function to update the weights in the network. This process of weights updating using the gradient of the cost function, which makes it possible to reduce the error in the subsequent iterations, is known as **backpropagation** [263, 264, 265, 266, 267, 268]. In backpropagation, the network motion is backward: the error along with the gradient flows back from the outer layer through the hidden layers and the network parameters are updated [262].

2.3.3.3.2 Convolutional neural networks specificity

In CNNs, the principle is the same as in NNs, with the exception of the term 'convolutional' and the following particularities. In general, CNN was shown to excel in a wide range of computer vision applications [265], such as face recognition, object detection, powering vision in robotics, self-driving cars, image processing tasks [269, 270] as well as sequential data analysis [191]. Its high performance and its easiness in training are two of the main factors driving the popularity of CNN over the

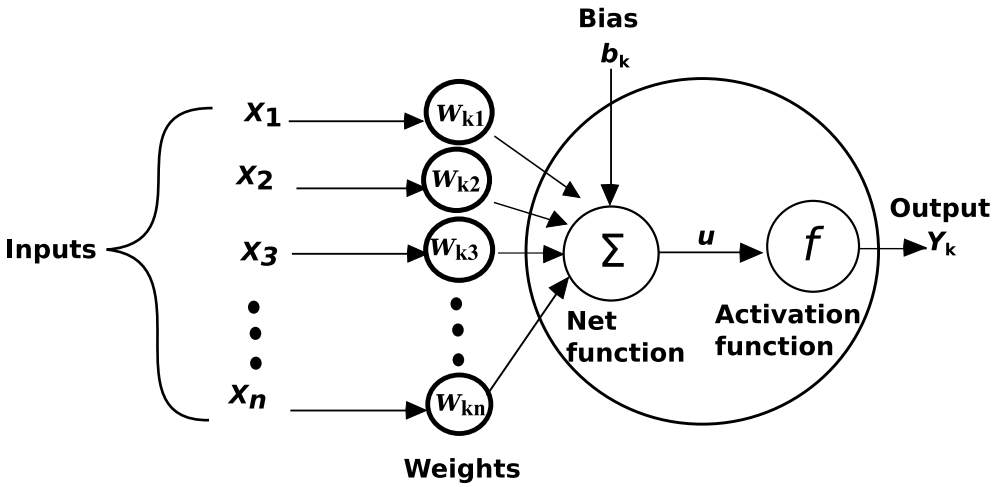

Figure 2.9: Structure of an artificial neuron. Adapted from [262].

last years [249]. Simply put and as shown in Figure 2.10, typical CNN architecture comprises two main parts:

- Convolutional base, which is composed by a stack of convolutional and pooling layers [271]. The main goal of the convolutional base is to extract features from the inputs and to reduce the spatial dimensions of the inputs for the next convolutional layer. In the convolutional layers, various kernels are used to convolve the inputs as well as the intermediate feature maps, generating various feature maps. Pooling layers are in charge of subsampling or downsampling, i.e., the operation consisting of width and height of the feature maps reduction. [261, 263, 264, 272, 273].

- Classifier, which is usually composed by fully connected layers. The main goal of the classifier is to classify the network inputs based on the detected features. In fact, the fully connected layers eventually interpret the feature representations arising from the succession of several convolutional and pooling layers stacked on top of each other and perform the function of high-level reasoning [238, 258, 274]. The derived output either could be fed forward into a certain number of categories for classification or could be considered as a feature for further processing.

One of the limitations of the classical CNN is that it only performs mono-output tasks.

The CNN used in this thesis has been adapted from the classical one to be able to fill this gap by achieving multi-output tasks useful for time series outputs.

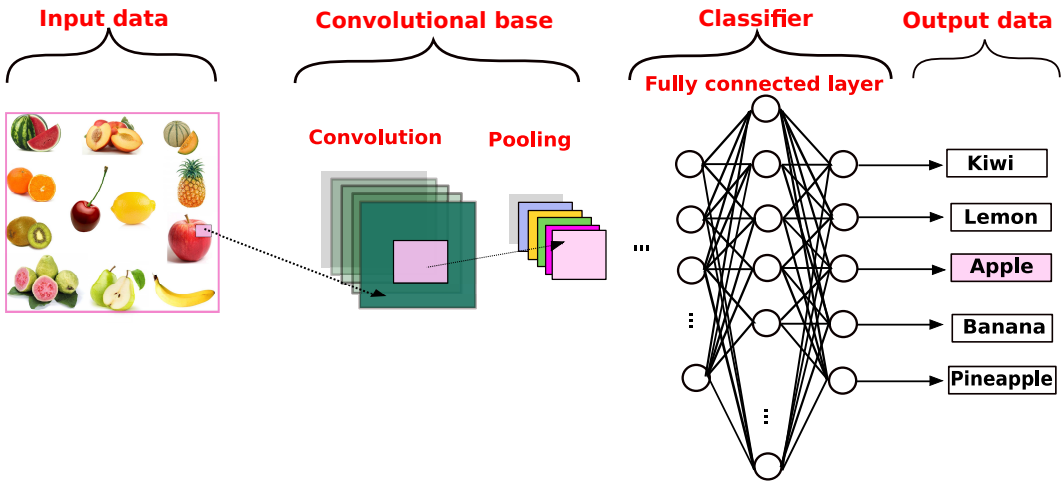

Figure 2.10: Typical CNN architecture for a fruit image classification task. An image is input directly to the network, and this is followed by several not fully connected stages of convolution and pooling. Then, representations from these operations feed one or more fully connected layers. Finally, the last fully connected layer outputs the class label.

Chapter key points:

- **An introduction to low-dimensional and high dimensional models of the CVS existing in the literature is presented.**
- **Particular emphasis is placed on the lumped-parameter models used in this work.**
- **Mathematical tools used in this work for model confidence evaluation and for model parameter estimation namely, uncertainty analysis, sensitivity analysis and ensemble Kalman filter, machine learning, respectively, are introduced.**

CHAPTER 3

General context, motivations and objectives

Chapter contents

After introducing in the two previous chapters the general concepts used in this thesis, it is now time to enter gradually into the heart of the matter with this third chapter. Closing the introductory part of this manuscript, this short chapter outlines the general context of the thesis as well as its objectives and contributions. The thesis outline is then presented while the main results of each chapter are provided.

3.1 General context and motivations

Cardiovascular diseases (CVDs) have become a public health concern nowadays [275]. As outlined in chapter 1, according to the World Health Organization (WHO), CVDs are the leading cause of death globally, taking an estimated 17.9 million lives each year, or 31 % of all global deaths [78, 276]. Among these diseases, cerebral aneurysms, which are present in 3.6 to 6% of the world's population, require special attention as they carry a high risk of mortality in case of rupture. Indeed, once bleeding or rupture of cerebral aneurysm occurs, blood spreads into the subarachnoid space around the brain leading to a subarachnoid hemorrhage (SAH) [1, 2], which quickly becomes life-threatening with a 40% risk of death [3, 4, 5].

One of the key factors identified to be associated with the formation and the risk of rupture of cerebral aneurysms is the blood pressure fluctuations in cerebral arteries [6, 7, 8, 9, 10, 11]. Therefore, a technique to estimate/assess these pressure variations in cerebral arteries, in a noninvasive way, would be beneficial and of great interest in clinical decision-making and treatment of cerebral aneurysms. In this context, Lal et al. have recently implemented an inversion platform based on EnKF coupled with a 0D reduced order blood flow model with the aim of estimating patient-specific blood pressure fluctuations in cerebral arteries [12]. This inversion procedure, which involves patient-specific cerebral Magnetic Resonance Angiography and Magnetic Resonance Imaging (MRA & MRI) acquisitions, uses a complex arterial network consisting of 33 arteries including the patient-specific cerebral circulation (circle of Willis). However, the accuracy and robustness of this platform with respect to some modeling assumptions and clinical data processing still need to be addressed and require therefore more investigation. Moreover, this inversion procedure, which requires inverse problem solving, is too long (no stopping criterion has been set there), expensive, difficult to integrate in a medical acquisition device, and its results still need to be validated. This begs the following questions:

- What is the impact of the different sources of uncertainty related to the models

used, the assumptions formulated as well as the patient-specific clinical data on the inversion procedure outcomes?

- Can the uncertainties on the procedure results be quantified and how? How reliable are these results?
- Is it possible to set a stopping criterion in the inversion procedure to avoid unnecessary over-solving and to reduce calculation time?
- How is it possible to accelerate this pressure estimation procedure, specifically in clinical situations where faster diagnosis is always requested?

Uncertainty analysis (UA) and sensitivity analysis (SA), on the one hand, and machine learning (ML), on the other hand, are two appropriate approaches to address these issues.

3.2 Thesis aims

As a continuation of the ongoing inversion procedure initiated by Lal et al. [12] for patient-specific cerebral blood pressure estimation, this thesis is first intended to scrutinize the accuracy and robustness of the inversion with respect to various sources of uncertainty related to the models, the formulated assumptions and clinical data used, and to set a stopping criterion for the procedure algorithm. To this aim, UA and several sensitivity analyses have been carried out. The second objective of this thesis is to illustrate how ML used together with a Convolutional Neural Network (CNN) can be a very good alternative to the inversion procedure for blood pressure estimation in real-time.

3.3 Thesis outline

To efficiently address the questions raised in Section 3.1 and be able to achieve the objectives set and enumerated in the previous Section (Section 3.2), this manuscript is

structured as follows:

3.3.1 Part I - General concepts and objectives

Composed of three chapters, this introductory section is intended to provide the reader with the basics necessary to understand this manuscript. It consists of the two previous chapters as well as the current one.

Chapter 1 - An introduction to the cardiovascular system

As already pointed out, this relatively short chapter is purely terminological, descriptive and explanatory in nature, and is intended to provide the reader and the non-specialist with a rudimentary knowledge of the human CVS and related diseases. The complexity of the CVS is just sketched because the main goal of this chapter is to introduce the bulk of the key terms related to the human CVS used in this manuscript.

Chapter 2 - A brief overview of current cardiovascular models and main mathematical tools.

As already highlighted, this second chapter of the first part introduces the reader to the models and mathematical tools used in this thesis. It begins with a literature review on the major existing physical models of the CVS with a focus on the 0D compartment models, which are the type of model used in this work. Then, it introduces the basics of the mathematical tools used in this work: the EnKF-based technique previously used in the inversion platform to estimate hemodynamic parameters is recalled while the basics of uncertainty analysis (UA), sensitivity analysis (SA), and machine learning (ML) approaches are introduced.

Chapter 3 - General context, motivations and objectives.

This is the chapter in progress. As mentioned at its beginning, the general context, the objectives, and the contributions of the thesis are set out in this chapter. Last

introductory chapter, it also provides the outline of the thesis, as well as the main results of each chapter.

3.3.2 Part II - Uncertainty and sensitivity analysis

Consisting of two chapters, this part attempts to achieve the thesis' first major objective. On the one hand, it examines the accuracy and robustness of the inversion procedure proposed by Lal et al. [12], with respect to various sources of uncertainty related to the models, the assumptions formulated and the clinical data used, and on the other hand, it sets a stopping criterion for the EnKF-based algorithm used in the inversion procedure.

Chapter 4 - Backward sensitivity analysis and reduced-order covariance estimation in noninvasive parameter identification for cerebral arteries.

In this chapter the accuracy and robustness of the inversion procedure previously developed by Lal et al. [12] for patient-specific noninvasive cerebral blood pressure estimation is scrutinized. A brief summary of the physical model adopted, the inverse problem solved, patient-specific clinical data and typical results previously obtained is first recalled. Then, the issue of the accuracy and robustness of the inversion procedure is discussed in detail by means of uncertainty quantification and sensitivity analysis techniques with respect to different sources of uncertainty.

This chapter also proposes a deterministic construction for backward uncertainty propagation allowing a low-complexity estimation (just as reliable as estimation using ensemble approach) of the covariance matrix of the patient-specific arterial network hemodynamic parameters after inversion.

This chapter has been published as: Rapadamnaba, R, Nicoud, F, Mohammadi, B. "Backward sensitivity analysis and reduced-order covariance estimation in noninvasive parameter identification for cerebral arteries" in *International Journal for Numerical*

Methods in Biomedical Engineering, 2019; 35:e3170. https://doi.org/10.1002/cnm.3170 [277].

Chapter 5 - Global sensitivity analysis for assessing the parameters importance and setting a stopping criterion in a biomedical inverse problem.

Chapter 5 provides an extension of uncertainty and sensitivity analysis performed in chapter 4. In concrete terms, it shows how to obtain, in addition to the standard deviations available after the inversion procedure, an apportioning of the total uncertainty in the outputs of the patient-specific blood flow model (i.e., in the inversion results) into small portions of uncertainty due to model input parameters. Some statistical indicators for analyzing model parameters, such as the Sobol' indices, are computed in order to identify both the importance rank of the patient-specific model input parameters and the influence of the interactions between these parameters on the output variance of the inversion procedure. In the light of the conclusions drawn from this SA, a suitable stopping criterion is proposed for the inversion algorithm with very limited extra calculation avoiding unnecessary over-solving and therefore significantly reducing the calculation time to solution.

This chapter has been submitted for publication as: Rapadamnaba, R, Ribatet, M, Mohammadi, B. "Global sensitivity analysis for assessing the parameters importance and setting a stopping criterion in a biomedical inverse problem" in *International Journal for Numerical Methods in Biomedical Engineering*, 2020 [278].

3.3.3 Part III - Machine learning for cerebral pressure

This last part of the manuscript is intended to achieve the second major objective of this thesis. Containing only one chapter, it shows how machine learning used together with CNNs, can be a very good alternative to the time-consuming and costly inversion

procedure previously implemented by Lal et al. [12] for assessing cerebral arterial pressure.

Chapter 6 - Augmented patient-specific functional medical imaging by implicit manifold learning

This chapter uses machine learning together with an implemented CNN, referred to as implicit CNN, as an alternative to the EnKF-based inversion procedure for patient-specific blood pressure estimation. Using a synthetic database organized according to the availability of patient-specific data from medical images and connecting the geometrical parameters and mechanical characteristics of the arteries to the blood flow rates and pressures in the patient-specific arterial network, a CNN is built and trained. Once properly trained, the resulting neural network is used in order to predict blood pressure in cerebral arteries noninvasively in nearly real-time. Finally, to validate the results, prediction of the patient-specific blood pressure using the inversion procedure and that using the implicit CNN are compared.

This chapter has been published as: Rapadamnaba, R, Nicoud, F, Mohammadi, B, "Augmented patient-specific functional medical imaging by implicit manifold learning" in *International Journal for Numerical Methods in Biomedical Engineering*, 2020; 36:e3325. https://doi.org/10.1002/cnm.3325 [278]

Material from this chapter has also been presented as "Estimation non invasive de la pression dans les artères cérébrales" at 9ème Biennale Française des Mathématiques Appliquées et Industrielles, 13 - 17 mai 2019, Guidel Plages (Morbihan), France.

Conclusion

Chapter 7 - Main results, discussions and perspectives

In this last chapter, the main findings and conclusions drawn from this thesis are recalled. Then, the physical modeling choices, the assumptions formulated and their consequences are discussed. Finally, possible directions and perspectives for future research are proposed.

Chapter key points:

- **The need to contribute to the fight against cerebral aneurysms and to continue the research work already initiated in this direction through the estimation of patient-specific cerebral pressure motivates this work.**
- **The aim is to assess the reliability and robustness of the available cerebral pressure estimation procedure and to propose a competing alternative method.**
- **The manuscript consists of 7 chapters arranged in 3 different parts.**

Part II

Uncertainty and sensitivity analysis

CHAPTER 4

Backward sensitivity analysis and reduced-order covariance estimation in noninvasive parameter identification for cerebral arteries

Chapter contents

This chapter has been published as:

Rapadamnaba, R, Nicoud, F, Mohammadi, B. "Backward sensitivity analysis and reduced-order covariance estimation in noninvasive parameter identification for cerebral arteries" in *International Journal for Numerical Methods in Biomedical Engineering*, 2019; 35:e3170. https://doi.org/10.1002/cnm.3170 [277].

The numbering of sections, figures and tables in the original manuscript have been altered for this chapter, to be consistent with the thesis chapter numbering. The references have been incorporated into the global references at the end of the thesis.

The rest of this chapter presents sensitivity analysis of the inversion procedure previously developed for blood pressure estimation with respect to the inlet flow rates, the choice of boundary conditions and the symmetry assumption in the network terminations. The aim of this chapter is to examine the robustness and the reliability of the inversion procedure by quantifying the uncertainties in the inversion outcomes and propose an alternative backward uncertainty quantification construction to the inversion procedure. An interesting result of this analysis is that uncertainties in the inversion outcomes are of the same order of magnitude as the uncertainties in the model input parameters.

Abstract

Using a previously developed inversion platform for functional cerebral medical imaging with ensemble Kalman filters, this work analyzes the sensitivity of the results with respect to different parameters entering the physical model and inversion procedure, such as the inlet flow rate from the heart, the choice of the boundary conditions, and the nonsymmetry in the network terminations. It also proposes an alternative low complexity construction for the covariance matrix of the hemodynamic parameters of a network of arteries including the circle of Willis. The platform takes as input patient-specific blood flow rates extracted from magnetic resonance angiography and magnetic resonance imaging (`dicom` files) and is applied to several available patients data. The paper presents full analysis of the results for one of these patients, including a sensitivity study with respect to the proximal and distal boundary conditions. The results notably show that the uncertainties on the inlet flow rate led to uncertainties of the same order of magnitude in the estimated parameters (blood pressure and elastic parameters) and that three-lumped parameters boundary conditions are necessary for a correct retrieval of the target signals.

Keywords: covariance matrix, ensemble Kalman filter, parameter estimation, reduced order compartment blood model, uncertainty quantification, sensitivity analysis.

4.1 Introduction

Cardiovascular diseases are obviously major health concerns nowadays as according to the World Health Organization, they are the leading cause of death globally [279]. Among these diseases, intracranial aneurysms, which usually take place in the circle of Willis (coW) [61], require special attention. Indeed, according to the Brain Aneurysm Foundation, the cerebral aneurysm is known as silent killer because in the most of the cases, it is completely asymptomatic and is discovered by chance frequently in people

undergoing brain imaging, such as magnetic resonance angiography and magnetic resonance imaging (MRA & MRI), and for other reasons, such as evaluation of headaches, after head trauma, or in work-up of other neurological symptoms [280]. Often in these situations, the aneurysm itself is an incidental finding, unrelated to the symptoms that prompted the imaging, but a ruptured aneurysm can cause life-threatening blood loss, which leads to death. That is why over the last few decades, to better understand and identify the mechanisms linked to aneurysm formation and rupture, many researchers, across a series of studies, made every effort to establish the key factors that contribute to the development and the rupture of cerebral aneurysms [6, 7, 8, 11, 281, 282, 283]. Some addressed the substantial challenges facing them regarding hemodynamics and the blood flow mechanism in the coW [284, 285, 286, 287]. Their main objective was to understand the factors increasing the risk of stroke and the blood flow distribution in the brain. Others have demonstrated the importance of fluid-structure interaction in patient-specific analysis of cerebral aneurysms [288, 289, 290, 291] and also the considerable influence of patient-specific anatomy of cerebral arterial network on blood flow patterns in local cerebral aneurysms [284, 292, 293, 294, 295, 296, 297, 298].

In previous studies [299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309], still others have developed an algorithm based on a data assimilation technique to estimate hemodynamic parameters such as elastic properties of arteries, arterial compliance, and boundary condition parameters (Windkessel boundary parameters and reflection coefficients).

In our recent publications [12, 153], the usefulness of this kind of algorithm in estimating the hemodynamic parameters (e.g., the wall thickness and the Young modulus), which are difficult to identify noninvasively, was demonstrated. And to examine the robustness of the approach, several sensitivity analyses have been carried out. More precisely, the behavior of the algorithm has been analyzed for

- different initial guesses for the parameters,

- different levels of observations uncertainty,
- the effect of bias in known parameter values and the type of observations,
- the effect of the ensemble size q_{ens} on the estimated parameters,
- the location of observations and the number of observations n_{obs}.

This work considers other sources of uncertainties and studies their impacts on the inversion outcomes. More precisely, it will address

- the uncertainties on the inlet flow rate time series from MRA & MRI, which has been considered as deterministic in our previous works;
- the uncertainties related to the choice of the boundary conditions at the terminations;
- the uncertainties related to the symmetry assumption previously used between the left and right network terminations.

Another proposal in this paper is a deterministic construction for backward uncertainty propagation not using an ensemble approach and an alternative estimation of the covariance matrix of the estimated parameters after inversion. This construction can be used even without an ensemble formulation and with deterministic gradient-based minimization algorithms.

The general minimization framework in which the fluid-structure coupling is cast is recalled in Section 4.2. The cardiovascular network model, which is made of 0D lumped compartments, is presented together with the associated minimization problem in Section 4.3. Section 4.4 presents the set up, which will be used to illustrate the different results. This is from one of the patient-specific data provided by Montpellier University Hospital, *Gui de Chauliac*. Section 4.5 gathers the results of various sensitivity analyses. The inversion procedure is based on the ensemble Kalman filter (EnKF) algorithm described in Lal et al. [153]. These ensemble algorithms give

an estimation of the covariance matrix $\mathbf{cov_x}$ of the optimization parameters. This is compared to an alternative derivation of the covariance matrix in Section 4.6. Section 4.7 discusses some limitations of the paper in the light of the model assumptions.

4.2 General settings

This section presents materials for this work. It follows in broad outline the description made in Mohammadi [310].

This study focuses on a class of minimization problems written as follows: where $\mathbf{x}$, $\mathbf{y_{obs}}$, and $\mathbf{z}$ are independent variables. Only the state variable, $\mathbf{y}$, depends on the cost function, J. The optimization parameter $\mathbf{x}$ belongs to $\mathbf{O}_{ad}$ the optimization admissible domain [311]. The physical meaning of all the variables will be given in Section 4.3.2.

This is a very general context, and it is important to address the effects of the variability in $\mathbf{z}$ and $\mathbf{y_{obs}}$ on the solution of the minimization problem. To analyze these, different approaches will be adopted: an ensemble approach (i.e., EnKF here) for the effect of the variability in $\mathbf{y_{obs}}$ and a separated treatment through a consideration of adequate ensemble of ensembles for the effect of the variability in $\mathbf{z}$.

To be more precise, attention is paid on functional J of the form

$$J(\mathbf{y}(\mathbf{x},\mathbf{z}),\mathbf{y_{obs}}) = \|\mathbf{y}(\mathbf{x},\mathbf{z}) - \mathbf{y_{obs}}\|_*, \tag{4.1}$$

where the state $\mathbf{y}(\mathbf{x},\mathbf{z})$ is solution of a state equation $F(\mathbf{y}(\mathbf{x},\mathbf{z})) = 0$, $\mathbf{y_{obs}}$ comes from a direct observation (measurements) of the system, and $\|.\|_*$ is a suitable norm. These will be specified in Section 4.3.2. The variable $\mathbf{y_{obs}}$ is assumed uncertain and independent, and its components are given by their probability density functions here assumed Gaussian $\mathcal{N}(\mu_i, \sigma_i^2)$, $\quad i = 1, ..., n_{obs}$ with means μ_i and variances σ_i^2.

4.3 Modeling and problem specification

4.3.1 Physical model

Many sophisticated and complex physical models representing the human cardiovascular system exist in the literature [284, 312, 313, 314, 315], the simplest of these being the 0D model (also called compartement model or lumped parameter model). It is the physical model adopted for this work. In this model, built using an electrical analogy [121, 305, 315, 316], the arterial network is divided into different compartments, each with a resistor (resistance R of blood due to blood viscosity and the vessel diameter), an inductor (blood inertia L), and a capacitor (compliance C of the artery corresponding to the quality of the latter to accumulate and release blood due to elastic deformations) as shown in Figure 4.1.

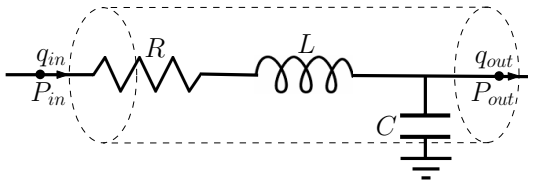

Figure 4.1: A single compartment circuit illustration.

The governing system of equations relating the variables R, L, and C for this model results from the Kirchhoff current and voltage laws (corresponding to the momentum and mass conservation principles here) applied to a single compartment assumed being filled with an incompressible Newtonian fluid. They take the form [121, 315, 316]

$$\begin{aligned} \frac{dP_{out}}{dt} &= \frac{q_{in} - q_{out}}{C} \\ \frac{dq_{in}}{dt} &= \frac{P_{in} - P_{out} - Rq_{in}}{L}, \end{aligned} \tag{4.2}$$

where for each artery, P_{in}, P_{out}, q_{in}, and q_{out} are inlet blood pressure, outlet blood pressure, inlet flow rate, and outlet flow rate, respectively. For arteries with a radius $r < 0.2$ cm, the inertial effect is ignored [317], and in this case, the flow rate is given by $q_{in} = (P_{in} - P_{out})/R$.

To calculate the parameters R, L, and C of each compartment, the following equations [305] are used : Hagen-Poiseuille equation for resistance, $R = 8\mu l/\pi r^4$, $L = \rho l/\pi r^2$, and $C = 3\pi r^3 l/2Eh$, where r, l, μ, ρ, h, and E are the radius of the artery, the length of the arterial segment, the blood viscosity, the blood density, the arterial wall thickness, and the Young modulus, respectively.

Each arterial segment of the full network including the coW is represented by a reduced-order 0D model containing the three elements R, L, and C (see Figures 4.3C,D, where each of the arterial segments is modelled with a single compartment). In this way, the full network is represented by a distributed lumped parameter model in which multiple lumped compartments are interconnected.

At the bifurcations, the enforcement of the mass conservation principle and of the continuity equation for pressure permits to prescribe the boundary conditions. In order to include the effect of the downstream vasculature, the blood flow model is coupled to the three-element Windkessel model (WK3-lumped parameter model) [286, 318] at the outlet of each terminal compartment. In the WK3 model, the equation relating the instantaneous blood pressure and the flow rate reads as follows:

$$\frac{dp(t)}{dt} + \frac{p}{R_D C} = R_P \frac{dq(t)}{dt} + \frac{qR_T}{R_D C}, \tag{4.3}$$

where p, q, C, R_P, and R_D are the instantaneous pressure at the inlet of the WK3 model, the instantaneous flow rate, the compliance, the proximal resistance, and the distal resistance of the vascular beds, respectively. $R_T = R_P + R_D$ expresses the total peripheral resistance. The sensitivity of the present approach with respect to the choice of the boundary conditions will be discussed in Section 4.5.3.

It is not easy to solve the first order differential equations 4.2 and 4.3 in view of the stiffness of the system. That is why the Fortran version of an implicit numerical integration solver `DVODE` [319, 320], available on `http://www.radford.edu/~thompson/vodef90web/`, is used to solve the system.

4.3.2 Problem Specification

With the physical model above adopted, the general optimization problem presented in Section 4.2 can be more specified. The optimization parameter $\mathbf{x}$ of size n are the hemodynamic parameters for each segment plus the number of parameters in the termination boundary conditions. The vector $\mathbf{z}$ of size m contains the parameters used to describe the flow rate from the heart. Thus, m corresponds either to the number of points in a discrete representation or to the number of parameters for a reduced-order representation of the signal as shown in Section 4.5.2. The variable $\mathbf{y_{obs}}$ (called observations) is flow rate time series from MRA & MRI acquisitions at given arteries as described in Section 4.4. With time-series observations $\mathbf{y_{obs}}(t)$, different functional J can be considered. Following Lal and al. [12, 153], this work aims at minimizing a time-dependent functional based on instantaneous incoming information:

$$J(t, \mathbf{y}(\mathbf{x}, \mathbf{z}, t), \mathbf{y_{obs}}(t)) = \|\mathbf{y}(\mathbf{x}, \mathbf{z}, t) - \mathbf{y_{obs}}(t)\|_* = \frac{1}{2}\|\mathbf{y}(\mathbf{x}, \mathbf{z},t) - \mathbf{y_{obs}}(t)\|^2.$$

4.4 Patient-specific clinical data

This section presents patient-specific clinical data used in the remainder of the paper. These data have been provided by the Department of Neuroradiology of the *Centre Hospitalier Régional Universitaire de Montpellier (CHRU), Montpellier, France*, and have been extensively described in Lal et al. [12] together with their acquisition procedures.

As a reminder, before and after image acquisition, arterial systolic brachial pressure (SBP) and diastolic brachial pressure (DBP) at rest were measured using a brachial automatic sphygmomanometer (Maglife, Schiller Medical). The pressure values measured were (115 and 72 mmHg) in the left brachial artery and (125 and 72 mmHg) in the right one. The ascending aorta and the internal carotid arteries (right and left ICAs) have been considered for the analysis of blood flow rates.

Figure 4.2 shows two pairs of the magnitude and phase contrast images acquired —

one for the ascending aorta (Figure 4.2B) and the other one for ICAs (Figure 4.2C) — and their corresponding blood flow rates (Figure 4.2A,C). These flow rates have been obtained from the GyroTools software, called GTFlow (https://www.gyrotools.com/gt/index.php/products/gtflow).

Figure 4.3 shows the typical 3D model (and morphology) of coW (see Figure 4.3A,B) determined through segmentation of a 3D time of flight magnetic resonance angiography (3D-TOF-MRA) of the patient's coW.

Using these different images, a complex arterial network of 33 arteries (Figure 4.3D) consisting of the aorta, vertebral, carotid, and brachial arteries together with an integral coW adapted from Alastruey et al. [284] has been modelled, and as shown in Table 4.1, some geometric measurements such as lengths and radii of coW's blood vessel have been obtained. More precisely, in the cerebral regions with best/high image quality, lengths and radii of cerebral arteries have been manually extracted from MRA using RadiANT DICOM Viewer software (http://www.radiantviewer.com/), while in the regions with poor image quality, they were obtained from average data reported in the literature. Also, other geometries, such as the carotid vascular tree one, could not be obtained because their acquisition requires the injection of contrast material called gadolinium, which is impossible to perform on healthy volunteers. To fill this kind of geometries and all the missing geometry of the full network, data reported in the literature were also necessary.

The inverse hemodynamic problem aims at identifying unknown parameters (the arterial stiffness and the WK3 model boundary parameters) for the network as shown in Figure 4.3D and as described in Lal et al. [12, 153]. In the parameter estimation problem, both available patient-specific flow rate waveforms for the right internal carotid (R-ICA; #21 in Table 4.1) and the left internal carotid (L-ICA; #23 in Table 4.1) were used as observations during EnKF assimilation steps. Blood rheological parameters μ and ρ were set at 0.004 Pa s and 1050 kg m^{-3} and at the inlet (ascending aorta, compartment #1 in Figure 4.3D), specific values of flow rates, q_{in}, measured by PC-

Table 4.1: Geometric parameters corresponding to arterial segments (and compartments) in Figure 4.3 measured from magnetic resonance imaging. The missing geometry (marked with an asterisk) of larger arteries is taken up from the average data in the literature [284, 286]. R indicates right and L, left.

Id	Name	l, cm	r, cm	Id	Name	l, cm	r, cm
1	Ascending aorta (AA)	4.00*	1.200*	18	L. post. comm. artery (L.PCoA)	1.20	0.075
2	Brachiocephalic (BraCe)	3.40*	0.620*	19	R. post. cerebral artery P2 (R.PCA, P2)	8.50	0.100
3	Aortic arch I (Aa I)	2.00*	1.120*	20	R. post. comm. artery (R.PCoA)	1.20	0.075
4	R. subclavian (R.Sub)	3.40*	0.423*	21	R. internal carotid I (R-ICA)	17.7*	0.200
5	R. common carotid (R.CC)	17.7*	0.250*	22	R. external carotid (R.ECA)	17.7*	0.150*
6	L. common carotid (L.CC)	20.8*	0.250*	23	L. internal carotid I (L-ICA)	17.7*	0.200
7	Aortic arch II (Aa II)	3.90*	1.070*	24	L. external carotid (L.ECA)	17.7*	0.150*
8	L. subclavian (L.Sub)	3.40*	0.423*	25	R. internal carotid II (R.ICA)	0.50	0.200
9	Thoracic aorta (ThorA)	15.6*	0.999*	26	L. internal carotid II (L.ICA)	0.50	0.200
10	R. brachial (R.BRA)	42.2*	0.403*	27	L. middle cerebral artery (L.MCA)	11.9	0.143
11	R. vertebral (R.VA)	14.8*	0.136*	28	L. anterior cerebral artery A1 (L.ACA, A1)	1.20	0.117
12	L. vertebral (L.VA)	14.8*	0.136*	29	R.middle cerebral artery (R.MCA)	11.9	0.143
13	L. brachial (L.BRA)	42.2*	0.403*	30	R. anterior cerebral artery A1 (R.ACA, A1)	1.20	0.117
14	Basilar (BAS)	2.70	0.150	31	R. anterior cerebral artery A2 (R.ACA, A2)	10.3	0.120
15	R. post. cerebral artery P1 (R.PCA, P1)	0.56	0.110	32	Anterior comm. artery (ACoA)	0.30	0.074
16	L. post. cerebral artery P1 (L.PCA, P1)	0.56	0.110	33	L. anterior cerebral artery A2 (L.ACA, A2)	10.3	0.120
17	L. post. cerebral artery P2 (L.PCA, P2)	8.50	0.100				

MRI (see Figure 4.2A,B) were imposed.

The results shown in the next section were obtained under the following assumptions on the unknown model parameters:

- The parameters R, L, C, and the WK3 boundary condition parameters are assumed to well reproduce the patient-specific description of the 0D blood flow model expressed by Equation 4.2.

- Eh, the product of Young modulus and thickness of arteries, is the unknown quantity to recover by data assimilation, and it is assumed to be given by this empirical formula [321]:

$$Eh = r(k_1\, e^{k_2 r} + k_3), \tag{4.4}$$

  where the radius (r) is measured from MRA. An estimation of the product Eh is found by looking for an estimation of the unknown constants k_i with their initial guesses as $k_1 = 2.0 \times 10^7\,\mathrm{g\ cm^{-1}s^{-2}}$, $k_2 = -22.0\,\mathrm{cm^{-1}}$ and $k_3 = 8.0 \times 10^5$ $\mathrm{g\ cm^{-1}s^{-2}}$.

- For each left and right pair of terminal compartments, the same WK3 parameters are assumed by symmetry. For instance, the terminal compartments #22 and #24 are assigned with the same WK3 boundary conditions. The parameters R_{P_i}, R_{D_i} and C_i where $i = 9, 10, 19, 22, 29, 31$ denotes the compartment numbers are also considered as unknown model parameters. Thus, 21 parameters consisting of six proximal resistances, six distal resistances, six compliances, and three constants defining the product Eh are estimated. The initial estimates for the proximal resistances and the compliances were taken from the data published by Alastruey et al. [284], and the initial guesses for the distal resistances R_D are chosen such that the ratio $R_P/R_T = 0.2$, i.e., $R_D = 4R_P$ [137].

4.5 Patient-specific results

This section presents typical convergence of the inversion algorithm and sensitivity analysis for the inverse problem with respect to several modelling issues and parameters of the direct model. Before presenting these results, it seems appropriate to clarify first the meaning of convergence in the remaining of this paper. More precisely, it is said that there is convergence if, on the basis of a visual inspection, the estimated parameters have only very minor variations from systole to diastole on the one hand, and on the other, from one cardiac cycle to the next. Quantitative parameters for convergence could be introduced, but it would weigh down the procedure.

4.5.1 Convergence

Figure 4.4 shows the target and model signals in time. One observes that there is a good agreement between the target and predicted flow rate waveforms after 10 cycles (about 8.35 s).

Figures 4.5 and 4.6 show the evolutions of the first two moments of the 21 estimated optimization parameters $\mathbf{x}$. One sees very different convergence patterns. Some variables do not fully converge even though the observations are well recovered.

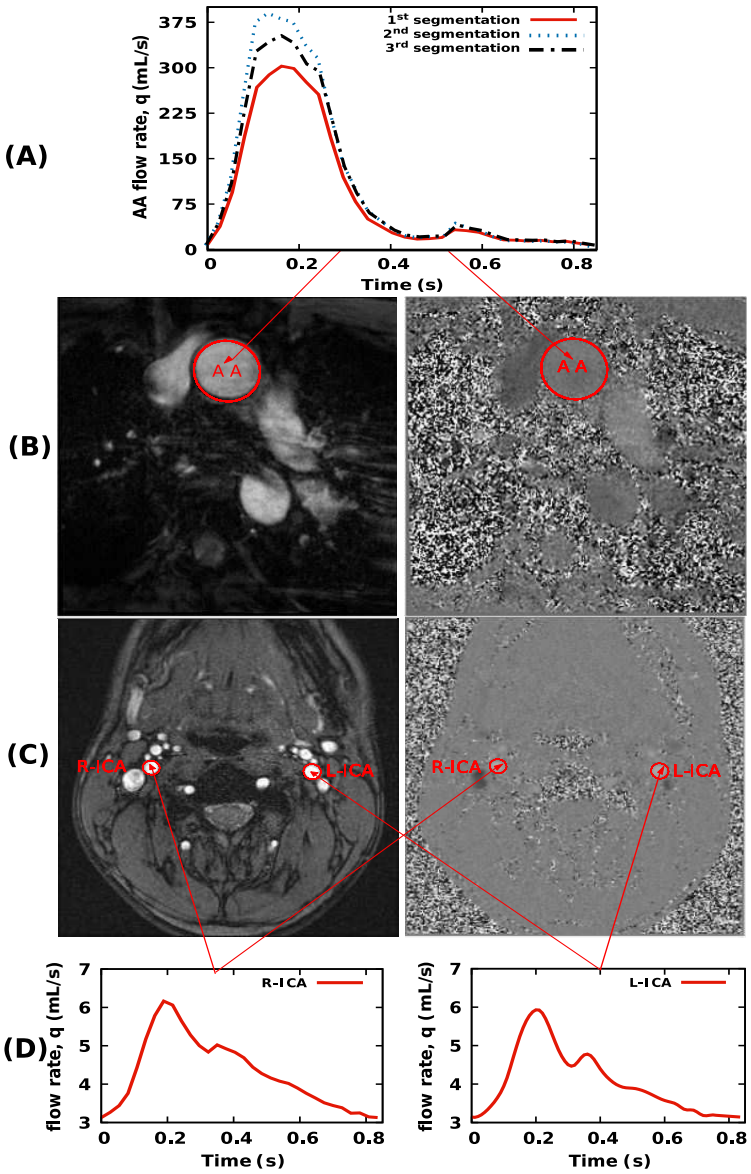

Figure 4.2: PC-MRI of the patient-specific ascending aorta and internal carotid arteries (right and left) showing the blood flow through one of the selected slices. In the center, on the left are the magnitude images and on the right are the phase contrast images B, and C, with Venc setting of 200 and 80 cm s^{-1} for the ascending aorta and the internal carotid arteries, respectively. The instantaneous blood flow rate values, $q(t)$ are acquired at each time frame and are plotted against time for one cardiac cycle as shown on the top for AA A, and at the bottom for ICAs D. In panel A, the signals obtained from three different segmentations performed by the same "operator" are reported to give a feeling of the uncertainty on the inlet flow rate.

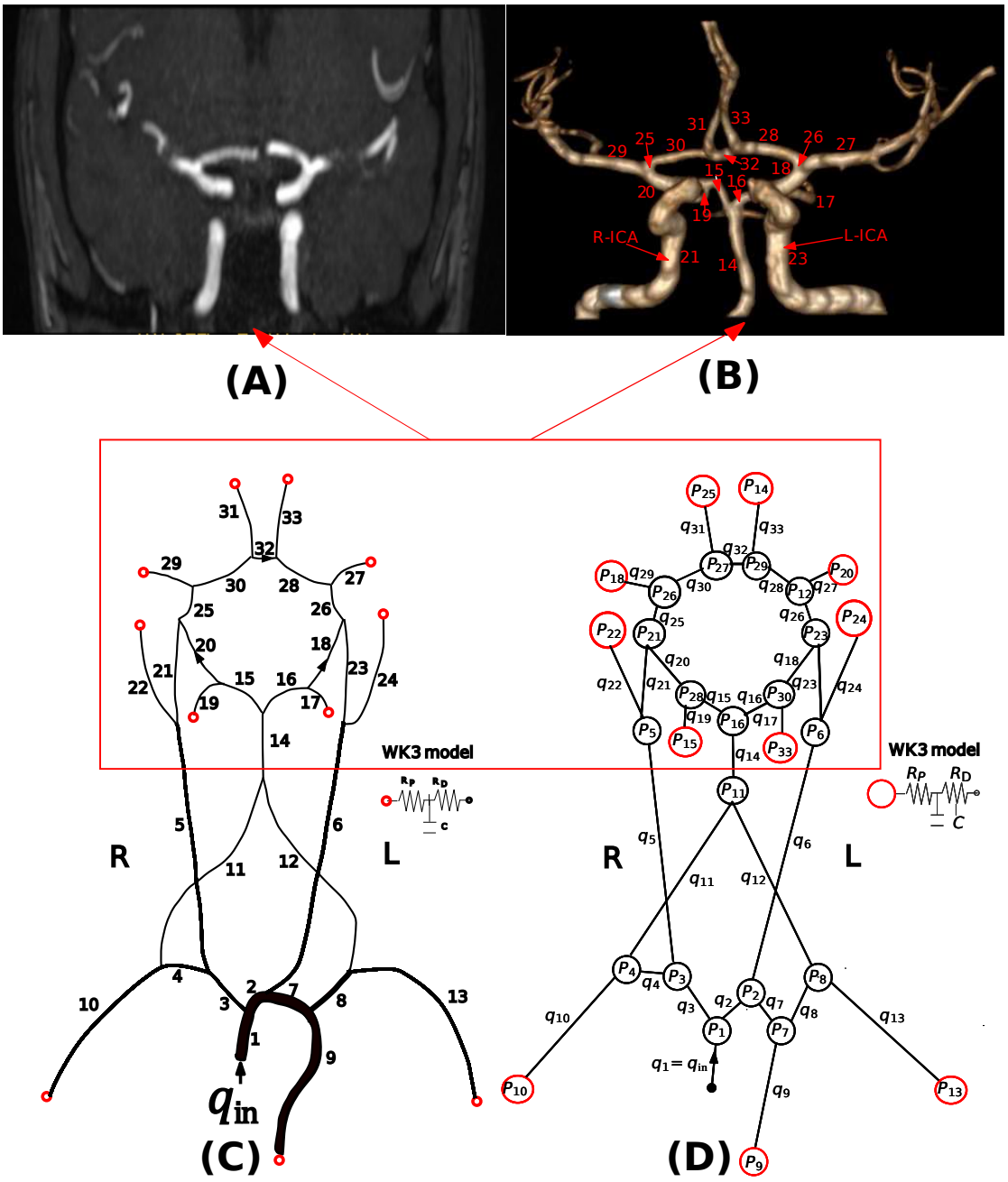

Figure 4.3: A TOF MRI scan A, and the corresponding segmented 3D model of the coW B, for the considered patient-specific case. The network [284] of a 1D blood flow model of the upper body arteries and of the coW C, and its equivalent compartment model D. The lines size indicates relative size of the arteries. The numbers on segmented model refer to the Ids of the arterial segments in Table 4.1. Arrows indicate the direction of flow. Flow rates are assigned the compartment numbers corresponding to those in Table 4.1. At the inlet (ascending aorta, compartment #1), specific values of flow rates, q_{in}, are imposed.

This suggests that those variables have small effects on the solution. The final estimates of the 21 parameters with their associated uncertainties are tabulated in Table 4.2. As the most important information is the elastic characteristics of the vessels, figure 4.7 shows the evolutions of three of the estimated Young moduli during the EnKF iterations. As the Young moduli are obtained through a same expression, they bare similar convergence patterns. The most important thing here to note is that all the Young moduli for the 33 arteries tend to realistic values with respect to the literature (i.e., between 0.1 and 1.2 MPa) [86, 284, 322].

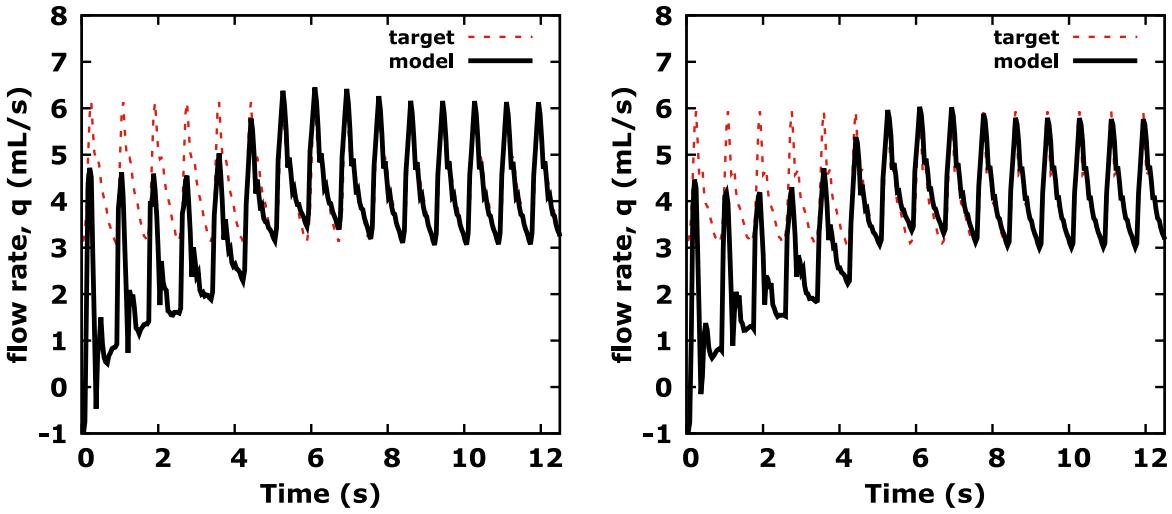

Figure 4.4: Comparison of the model simulated blood flow rate waveform in R-ICA (on the left) and L-ICA (on the right) to the target signals.

To illustrate this, a compared overview of some of estimated parameters with those reported in previous studies [137, 284, 286, 322, 323, 324] was carried out, and the model-simulated flow rates through all the termination models and carotid arteries were also examined. The results of the comparison of the parameters are shown in Figure 4.8 and the model-simulated flow one in Figures 4.4 and 4.9. These latter are also compared with blood flow rates values reported in Alastruey et al. [284] and Reymond et al. [286] and with clinically measured blood flow rates using MRI in Table 4.3. One notices that almost all estimated parameters were found to be in the same order of magnitude than those reported in other studies [137, 284, 286, 322, 323, 324]. Furthermore, one observes

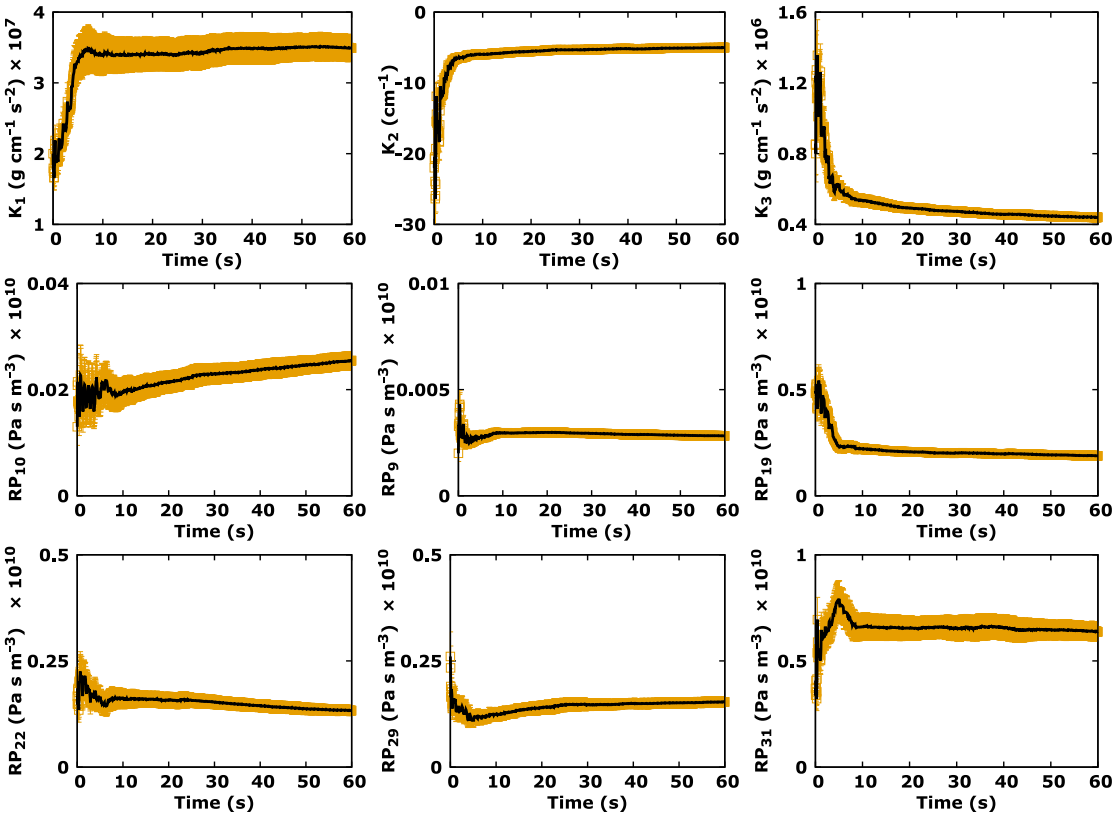

Figure 4.5: Histories of the mean and standard deviation of the nine first estimated parameters during EnKF iterations.

good overall agreement in flow waveforms but considerable differences in flow amplitude at all arterial terminations. All mean, systolic, and diastolic flow rates reported in Alastruey et al. [284] and Reymond et al. [286] are much higher than model outputs except those of MCA. This is to a large extent due to the difference of patients and models considered in investigated cases and especially to flow rate difference from the beginning in the ascending aorta (420, -33, and 96 mL s^{-1} for systolic, diastolic, and mean flow rate, respectively, in Reymond et al. [286] versus 302, 0, and 62.4 mL s^{-1} for systolic, diastolic, and mean flow rate, respectively, in the model). This prevents from properly comparing the different flow divisions of the model with those reported in Alastruey et al. [284] and Reymond et al. [286].

Despite these substantial differences, note that the model provided flow rate predictions that faithfully reproduced wave characteristics in the ICAs. This is

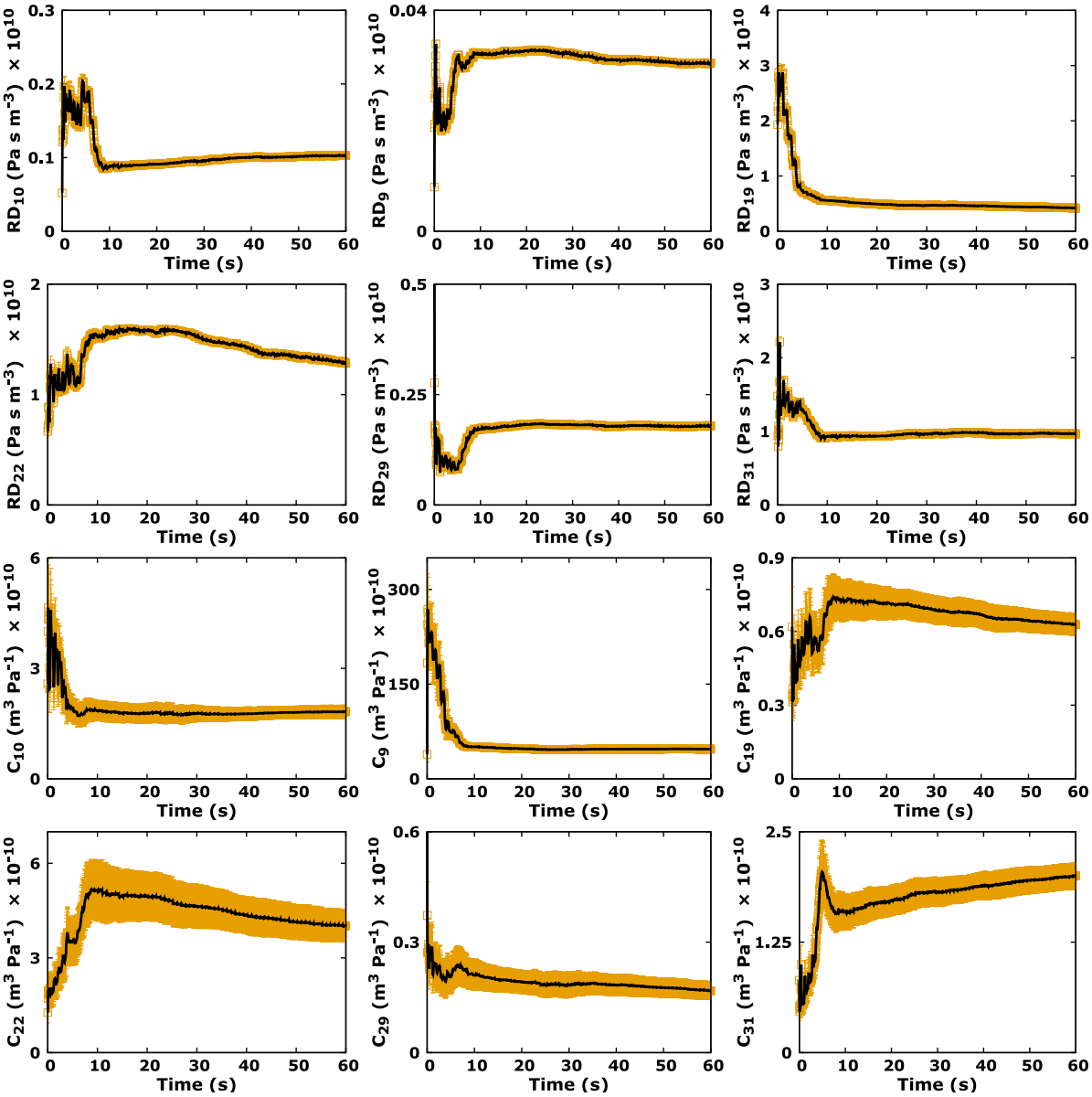

Figure 4.6: Histories of the mean and standard deviation of the 12 last estimated parameters during EnKF iterations.

Table 4.2: Estimated parameters with the associated errors for the patient-specific. The values of constants k_1, k_2, and k_3 are in $\times 10^7$ g cm^{-1}s^{-2}, cm^{-1}, and $\times 10^5$ g cm^{-1}s^{-2} respectively. The proximal (R_P) and distal (R_D) resistances are in $\times 10^9$ Pa s m^{-3} and the compliance (C) are in $\times 10^{-10}$ m^3 Pa^{-1}.

Parameter	Initial guess	EnKF estimate ± error
k_1	2.00	3.49 ± 0.19
k_2	-22.0	-5.00 ± 0.26
k_3	8.5	4.39 ± 0.24
R_{P9}	0.02	0.028 ± 0.001
R_{P10}	0.13	0.25 ± 0.017
R_{P19}	4.8	1.88 ± 0.15
R_{P22}	1.67	1.33 ± 0.12
R_{P29}	2.61	1.54 ± 0.11
R_{P31}	3.70	6.39 ± 0.43
R_{D9}	0.08	0.30 ± 0.0005
R_{D10}	0.52	1.03 ± 0.007
R_{D19}	19.32	4.19 ± 0.034
R_{D22}	6.68	12.88 ± 0.174
R_{D29}	10.44	1.79 ± 0.018
R_{D31}	14.80	9.69 ± 0.068
C_9	38.78	47.12 ± 2.62
C_{10}	2.58	1.82 ± 0.17
C_{19}	0.62	0.63 ± 0.043
C_{22}	1.27	4.02 ± 0.49
C_{29}	1.16	0.17 ± 0.023
C_{31}	0.82	2.00 ± 0.15

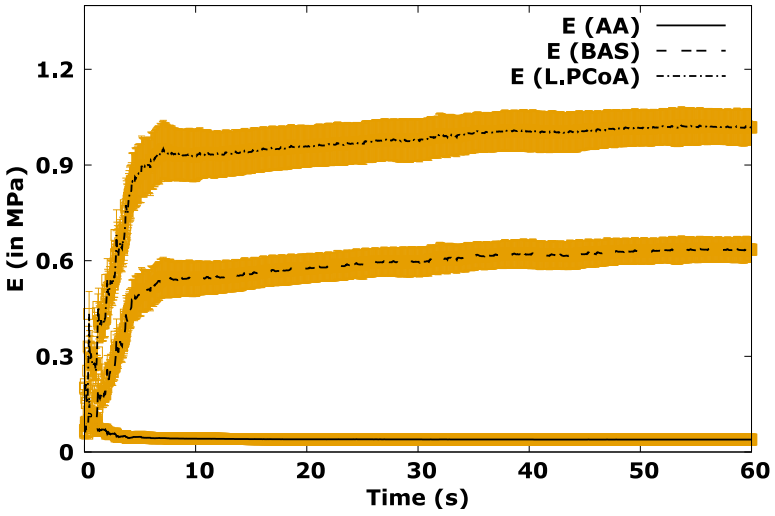

Figure 4.7: Mean and standard deviation of estimated Young moduli for three of the 33 segments of the network. The evolution is over 8.35 s corresponding to about 10 cardiac cycles.

illustrated in Figure 4.4, which shows the comparison between the target (clinically measured blood flow rates in the ICAs using MRI) and blood flow model simulations (predictions) based on the estimated parameters. From the results, the comparison between the assimilated 0D model and *in vivo* data (MRI) is fair. All mean, systolic, and diastolic flow rates measurements using MRI and model outputs in ICAs differ by less than 5%, and thus, one can conclude, in the absence of other clinically measured flow rates, that the 0D model considered here may very well predict blood flow rate in the entire arterial tree.

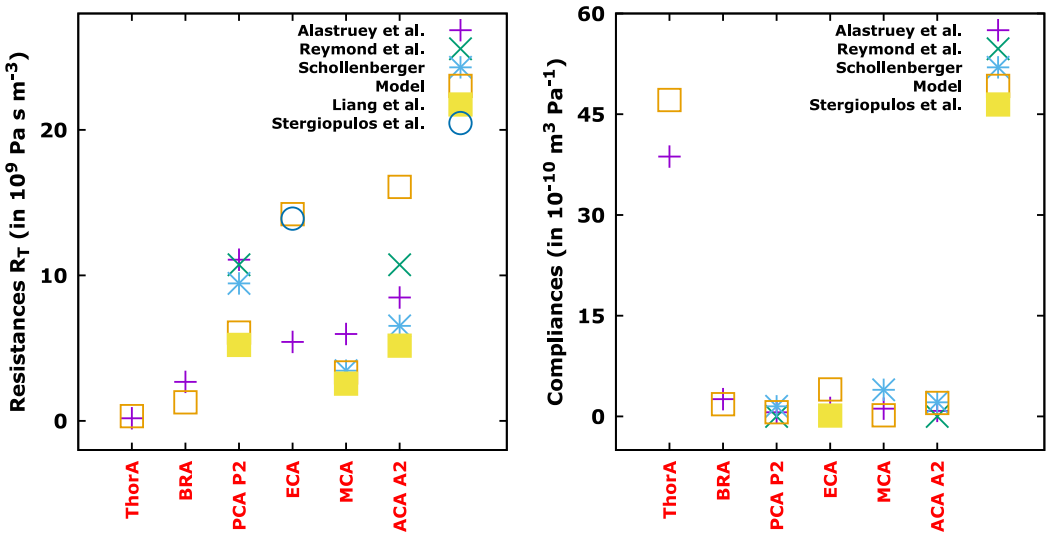

Figure 4.8: Comparison of estimated compliances and total peripheral resistances at the arterial terminations with reports in the literature. The estimated parameters using the assimilated compartment model (represented by empty squares) are in the same order of magnitude than those existing in the literature.

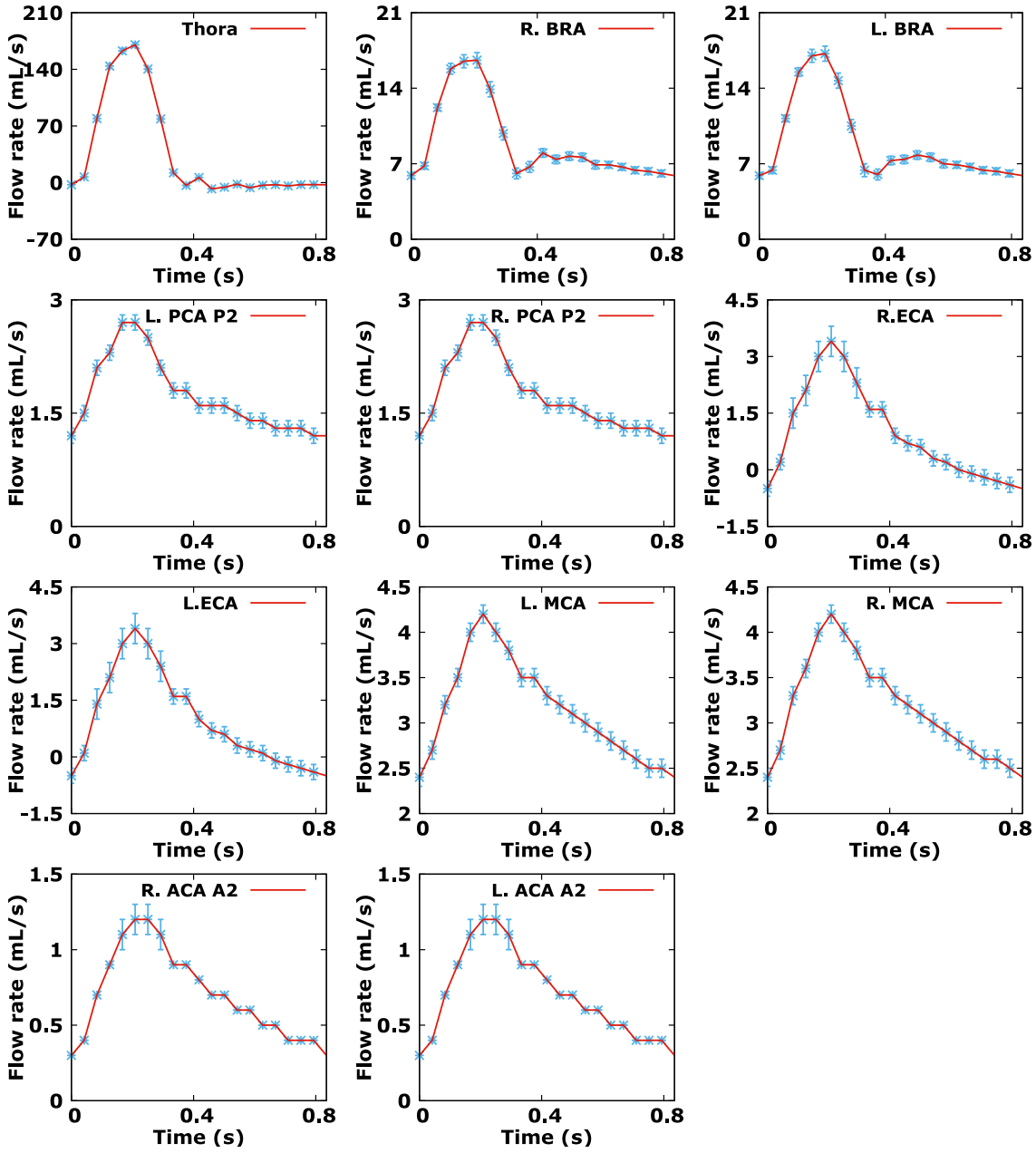

Figure 4.9: Mean and standard deviation of estimated blood flow rate at the arterial terminations.

Table 4.3: Quantification of blood flow at the arterial terminations, the CC and the ICAs: comparison of the cardiac cycle-averaged (mean), systolic and diastolic flow rates of model simulated values to measured values using MRI in the ICAs. The percentage difference between the model results and the MRI data are given in bold in parenthesis. The model results are also compared with the reported values in [284, 286].

	Systolic Flow Rate, mL/s			**Diastolic Flow Rate mL/s**			**Mean Flow Rate, mL/s**			
Artery Name	**Reymond et al. [286] in vivo/Model**	**MRI**	**Model Left/Right**	**Reymond et al. [286] in vivo/Model**	**MRI**	**Model Left/Right**	**Alastruey et al. [284] Left/Right**	**Reymond et al. [286] in vivo/Model**	**MRI**	**Model Left/Right**
AA	470/420	302.65	302	0.5/-33	7.98	0	–	103/96	87.82	62.4
Thora	303/235	–	170.4	-2.5/14	–	-7.7	–	68/71	–	37.91
			(±2.7)			(±1)				(±1.25)
BRA	–	–	17.2/16.6	–	–	5.9/5.9	–	–	–	9.02/9.02
			(±0.7)			(±0.3)				(±0.42)
PCA, P2	–	–	2.7/2.7	–	–	1.2/1.2	0.89/0.9	from	–	1.74/1.74
			(±0.1)			(±0.1)		0.88/0.85		(±0.1)
								to		
								2.08/2.08		
ECA	–	–	3.4/3.4	–	–	-0.5/-0.5	–	from	–	0.99/1
			(±0.2)			(±0.1)		1.5/1.5		(±0.13)
								to		
								2.5/2.7		
MCA	3.9/2.7	–	4.2/4.2	1.7/1.5	–	2.4/2.4	1.72/173	from	–	3.17/3.18
			(±0.1)			(±0.1)		1.8/2.1	–	(±0.1)
								to		
								2.78/2.78		
ACA, A2	–	–	1.2/1.2	–	–	0.3/0.3	1.16/1.16	from	–	0.72/0.72
			(±0.1)					1/1 to		(±002)
								1.25/1.47		
CC	22/22	–	9/9.4	2.5/0.7	–	2.4/2.5	–	6.5/5.4	–	5.12/5.27
			(±0.2)			(±0.1)				(±0.13)
ICA	9/9.8	5.93/6.13	5.79/6.15	2.2/2.7	3.09/3.13	3/3.06	4.06/4.31	3.6/3.9	4.15/4.29	4.15/4.29
			(±0.8/0.9)			(±0.8/0.9)				(±0.8/0.9)
			(2.3%/0.47%)			**(3.3%/2.3%)**				**(2.2%/0.3%)**

4.5.2 Sensitivity analysis with respect to the inlet flow rate $q_{in}(t)$

One important source of uncertainty in the inversion is due to the variability in the segmentation procedure by the "operator" in charge from PC-MRI data to define the variable called **z** in the minimization problem (5.3).

Let us consider $q_{in}(t)$ the patient-specific blood flow rate time series through one of the selected slices in the ascending aorta as shown in Figure 4.2A for one cardiac cycle. As presented in this figure, $q_{in}(t)$ is defined in a deterministic way. One however observes that different segmentations, even by a same "operator", lead to different flow rates time series. To analyze the sensitivity of the inversion with respect to this uncertainty, it is convenient to represent $q_{in}(t)$ using a reduced-order model. For that purpose, the model proposed in Alastruey et al. [284] involving two parameters which will be considered as stochastic here (see Figure 4.10) will be adapted as follows:

$$q_{in}(t) = \begin{cases} \alpha \sin(\pi t/\tau) & \text{if } t < \tau, \\ \varepsilon & \text{otherwise,} \end{cases} \tag{4.5}$$

where α (mL s^{-1}) $= \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$, τ (s)$= \mathcal{N}(\mu_\tau, \sigma_\tau^2)$, and ε (mL s^{-1}) $= \mathcal{N}(\mu_\varepsilon, \sigma_\varepsilon^2)$ are the components of **z** in (5.3). μ_α, μ_τ, and μ_ε are patient-specific. They are chosen solving a least square problem to fit the inflow rate shown in figure 4.2. This leads to the following values: $\mu_\alpha = 302$ mL s^{-1}, $\sigma_\alpha = 15$ mL s^{-1}, $\mu_\tau = 0.35$ s, $\sigma_\tau = 0.0175$ s, $\mu_\varepsilon = 21.65$ mL s^{-1}, and $\sigma_\varepsilon = 1.0825$ mL s^{-1}. The standard deviations that are of the order of 5 % of mean values model the uncertainty that an "operator" will introduce through the manual acquisition/segmentation steps from `dicom` format files.

To analyze the sensitivity of the inversion with respect to the uncertainty on $q_{in}(t)$, 30 inversions for Gaussian sampling in α, τ, and ε have been performed. An ensemble of size 30 has been used for all inversions. This ensemble size choice is based on a synthetic case study results that have shown that an ensemble of size 30 seemed to be good enough to estimate 21 model parameters for the prediction of cerebral arterial pressure with an error of less than 10 % [12]. To generate the 30 ensemble members, one proceeds as follows: from initial estimate of n unknown parameters (mean $\overline{x}_l$ and

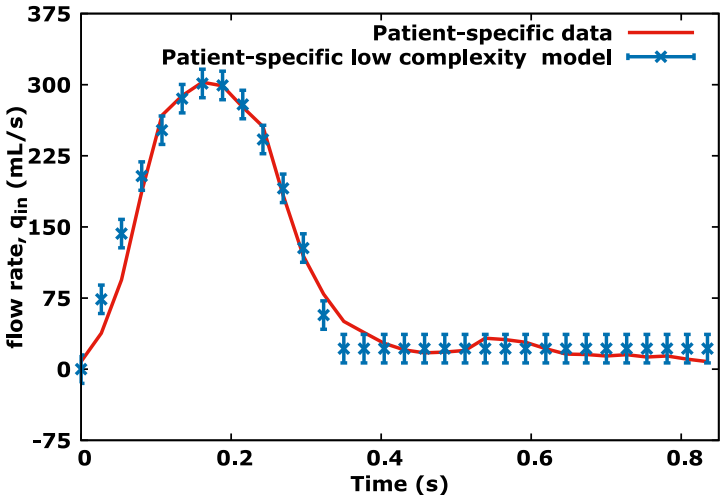

Figure 4.10: Uncertain inflow rate q_{in} and the low complexity model. The latter will be used for uncertainty propagation in inversion.

variance P_l for $l = 1, \ldots, n$), we randomly initialize an ensemble of parameters, $\mathbf{x}^i$, for $i = 1 \ldots, q_{\text{ens}}$ where $\mathbf{x}^i = (x_1, x_2, \ldots, x_n)$ and $x_l \sim \mathcal{N}(\overline{x}_l, P_l)$ for $l = 1, \ldots, n$. For more details, the reader is referred to Lal et al. [12]. It is interesting to illustrate the convergence not for the optimization variables $\mathbf{x}$ but for the Young moduli, which are obtained from following expression (4.4): $E = r(k_1\, e^{k_2 r} + k_3)/h$ where k_1, k_2, and k_3 are the three first components of $\mathbf{x}$.

Once all the inversions are completed, one can proceed with some statistics. Figure 4.11 shows the estimated mean value and standard deviation for the Young modulus of each segment. One sees that the "operator" uncertainty can have up to 5% impact on the elastic parameter estimation. This is therefore of the order of the patient-specific uncertainty in $\mathbf{z} = (\alpha, \tau, \varepsilon)$ in equation 4.5 and obtained assimilating uncertain patient data by the low complexity q_{in} model. This means that the procedure seems to work linearly in connection with the errors propagation. This is not an aberration as a result; on the contrary, this is satisfactory in that it reveals that the procedure does not seem to reduce nor to amplify the errors. However, q_{in} needs to receive very accurate treatment if one would like to have an uncertainty below say 5% on the outcomes, which is may be difficult to achieve given the nature of

the problem and also the way data are acquired.

Figure 4.12 presents the estimated mean and standard deviation for the right brachial pressure. It is quite reassuring to notice that despite the "operator" uncertainty, the predicted brachial pressure remains coherent with the auscultatory measure of the patient using a sphygmomanometer (i.e., between 72 and 125 mmHg) taken after the acquisition. This gives some indication for the level of uncertainty one can tolerate because of manual acquisition steps.

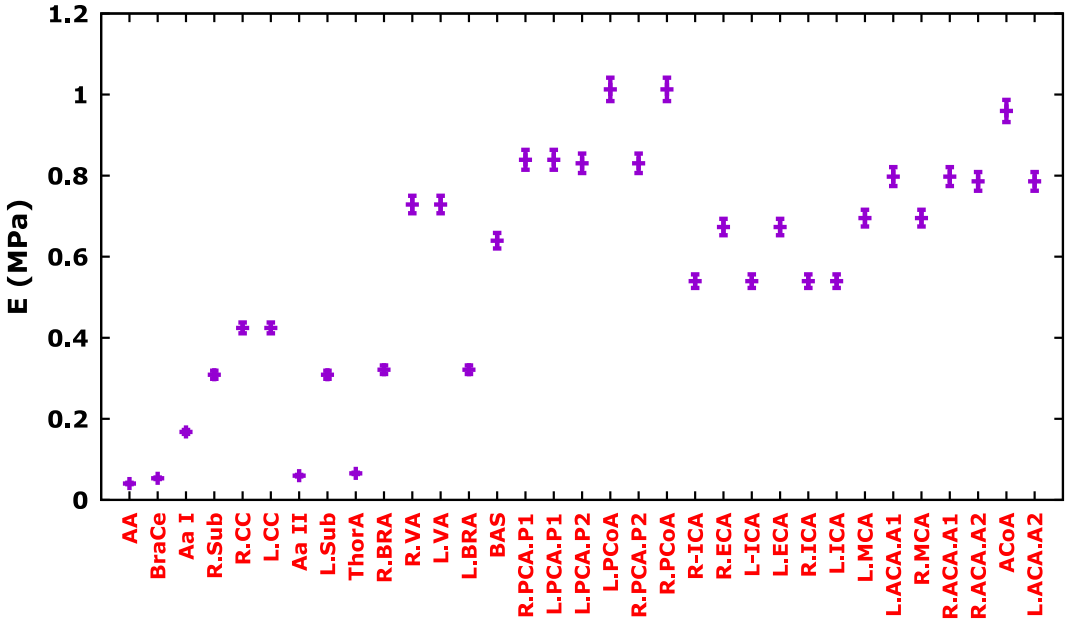

Figure 4.11: Mean and standard deviation of different segments Young moduli showing the impact of the uncertainty on the inlet flow rate $q_{in}(t)$ on the inversion.

4.5.3 Sensitivity analysis with respect to the boundary conditions

Another important source of uncertainty in the inversion is related to the choice of the boundary conditions at the outlet of each of the terminal compartments. As described in Section 4.3.1, the blood flow model uses a three-element Windkessel condition, which couples the instantaneous blood flow rate and pressure through equation 4.3. It would be interesting to see how the use of a simpler boundary condition will affect the inversion. More precisely, WK1 boundary conditions where the blood flow rates at terminations

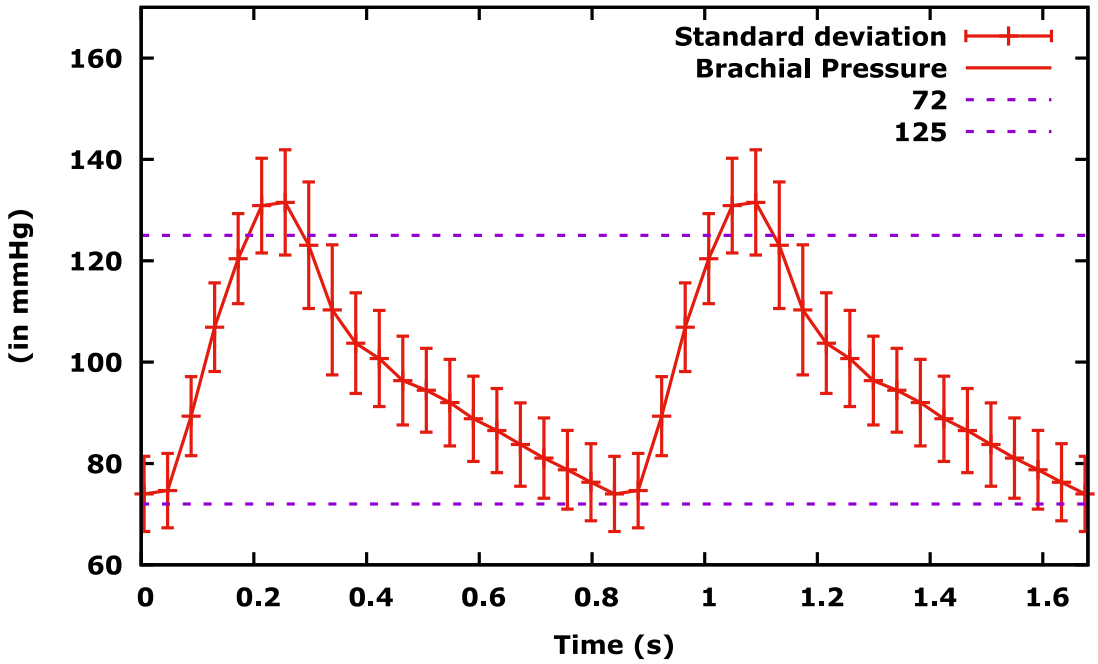

Figure 4.12: Right brachial pressure mean and standard deviation for two cardiac cycles showing the impact of the uncertainty on $q_{in}(t)$ on the inversion. The clinically measured SBP and DBP are 125 and 72, respectively.

are related to the instantaneous pressure through a simple algebraic relation involving the proximal and distal resistances are considered:

$$p = qR_T = q(R_P + R_D), \tag{4.6}$$

Hence, a WK1 condition is a WK3 one with zero compliance. One advantage of WK1 is a reduction of the number of optimization variables from 21 to nine: the six total resistances R_{T_i} with $i = 9, 10, 19, 22, 29, 31$ and the three constants k_1, k_2, and k_3 necessary for the definition of Eh. However, the same ensemble size of 30 is used for both inversions.

Figures 4.13, 4.14, and 4.15 show the impact on the inversion of a change of the boundary conditions from WK3 to WK1. This analysis shows that the WK1 conditions are clearly insufficient to describe the underlying physics of the problem. Indeed, despite the inversion's success in recovering the target flow rate at R-ICA and L-ICA segments, the brachial pressures are miscalculated. One also sees that the convergence of the

EnKF algorithm is not fully achieved. Actually, some parameters do not converge at all and this despite the ensemble size advantages in the WK1 inversion. Yet, to be adopted, the prescribed boundary conditions must allow at the same time to provide good convergence for the estimated parameters, to well reproduce both available patient-specific flow rate waveforms for the right internal carotid and the left internal carotid and to well predict brachial blood pressure. If one of these expected results is not achieved by the chosen boundary conditions, this choice cannot be relevant. As a conclusion, only WK3 conditions will be considered for the next sensitivity analysis regarding the interrogations of the assumed symmetry in the terminations boundary conditions.

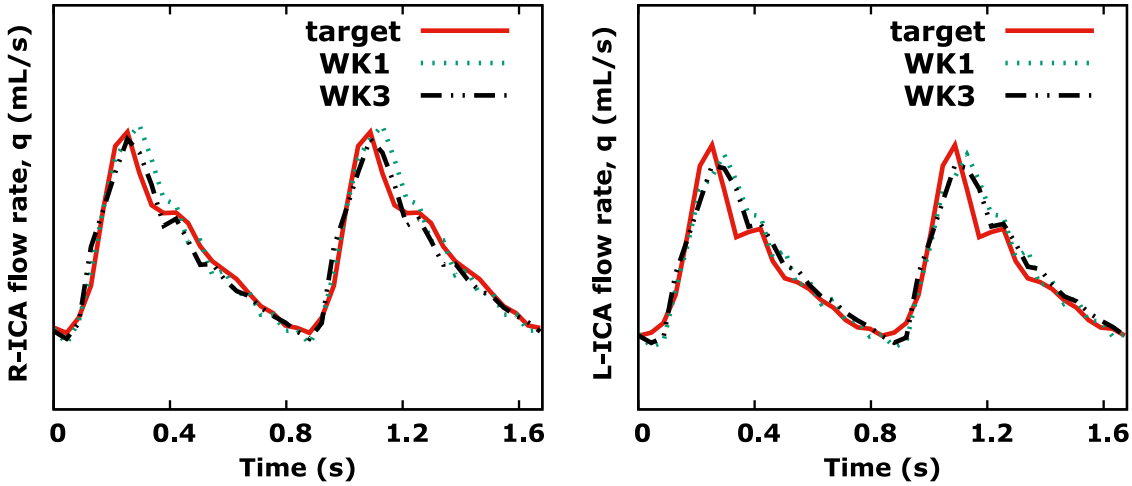

Figure 4.13: Flow rate waveform using WK1 and WK3 and comparison with the target flow rate in R-ICA and L-ICA.

As presented in Section 4.4, the same WK3 boundary conditions have been applied at the left and right pairs of the terminal compartments. To analyze the sensitivity of the inversion with respect to this assumption, one proceeds with several nonsymmetric scenarios. Unfortunately, to the knowledge of the authors, no indication on how the nonsymmetry takes place can be found in the literature. In other words, it is unclear if, for instance, the termination resistances or compliance are higher or lower on the left with respect to the right-hand side. Actually, it seems that the nonsymmetry is patient-dependent. The analysis will be carried out in three steps: nonsymmetric resistances

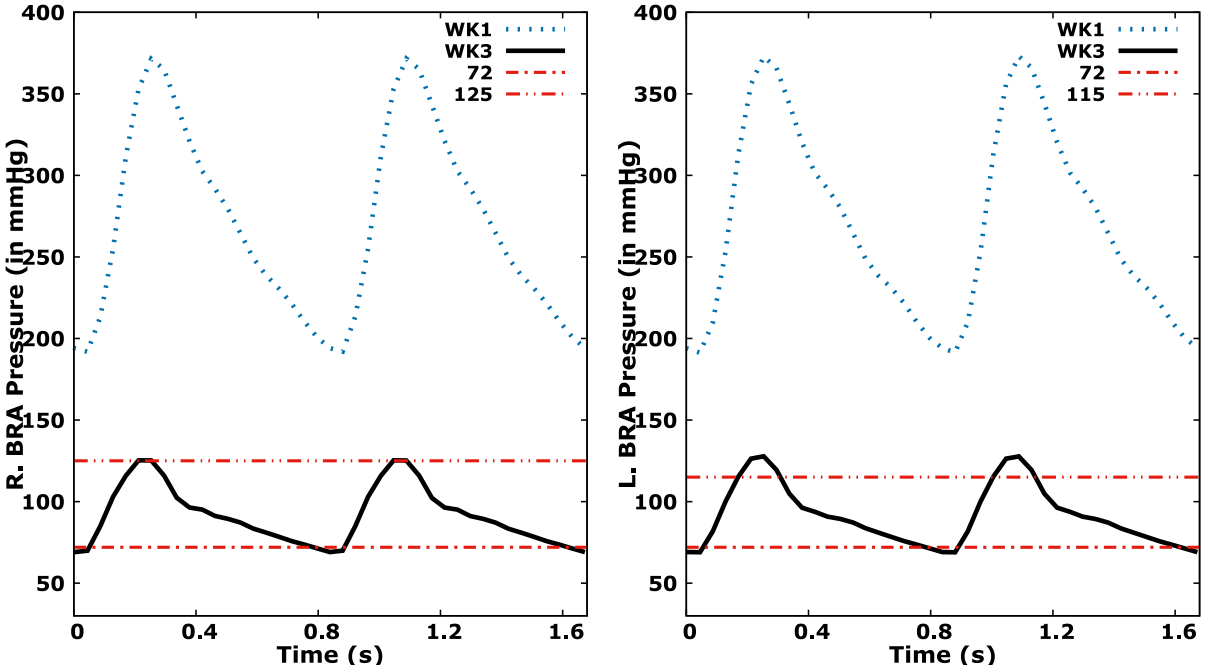

Figure 4.14: Comparison between R. and L. brachial pressures with WK1 and WK3. The clinically measured SBP and DBP in the right and the left brachial arteries are 125 and 72, and 115 and 72, respectively.

with symmetric compliances, nonsymmetric compliances with symmetric resistances, and nonsymmetric resistances and compliances.

Let us start analyzing the effect of nonsymmetry in the resistances on the inversion keeping the compliances symmetric. The nonsymmetry has been introduced considering $R_D^{left} = a_R R_D^{right}$ and $R_P^{left} = a_R R_P^{right}$ where a_R is a constant. Several inversions have been made for the values of $a_R = 0.5, 0.75, 1$ (symmetric WK3), 1.25 and 1.5, meaning that for each value chosen for a_R, the parameter identification procedure has been applied using the EnkF algorithm, which leads to the full set of physiological parameters. These values of a_R have been *a priori* chosen, and different sampling can be considered without difficulty. This is an illustration of how the inversion procedure can be used to analyze a possible hypothesis made by practitioners on the possibility of nonsymmetric behavior in terminations. Also, the sampling is made in order for the resistance to have admissible physiological values around the reference value "1" corresponding to the symmetric WK3 boundary conditions case. $R_D^{left,right}$ and $R_P^{left,right}$ replace R_D and

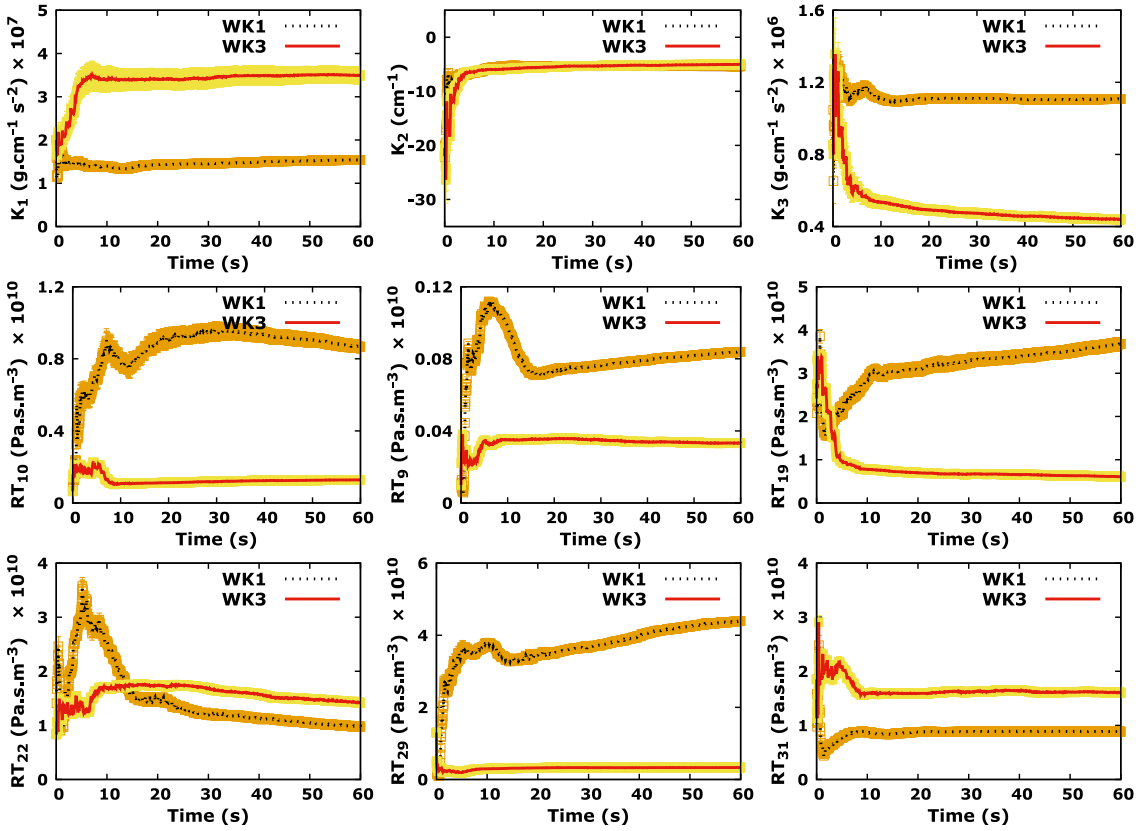

Figure 4.15: Convergence of estimated parameters during EnKF using WK1 and WK3. The shaded areas represent the standard deviation around the ensemble mean values (dashed and solid lines).

R_P in expression (4.3) in corresponding left and right terminations. Figures 4.16, 4.17, and 4.18 show the impact of this nonsymmetry on the inversion. One notices that a loss of symmetry on the resistances does not impact the flow rate. However, it plays a big role in pressure estimation as the mean brachial pressures and the SBP-DBP differences (`DeltaP`) increase with the level of resistances on the left terminations.

Now, let us see the impact of a nonsymmetry on the compliances for symmetric resistances. The nonsymmetry is again introduced through a multiplication by a constant a_C on the left compliances, $C^{left} = a_C C^{right}$. Again, several inversions have been performed for the values of $a_C = 0.5, 0.75, 1$ (symmetric WK3), 1.25 and 1.5. $C^{left,right}$ replace C in expression (4.3) in corresponding left and right terminations. Figures 4.19, 4.20, and 4.21 show the impact of this nonsymmetry on the inversion. It

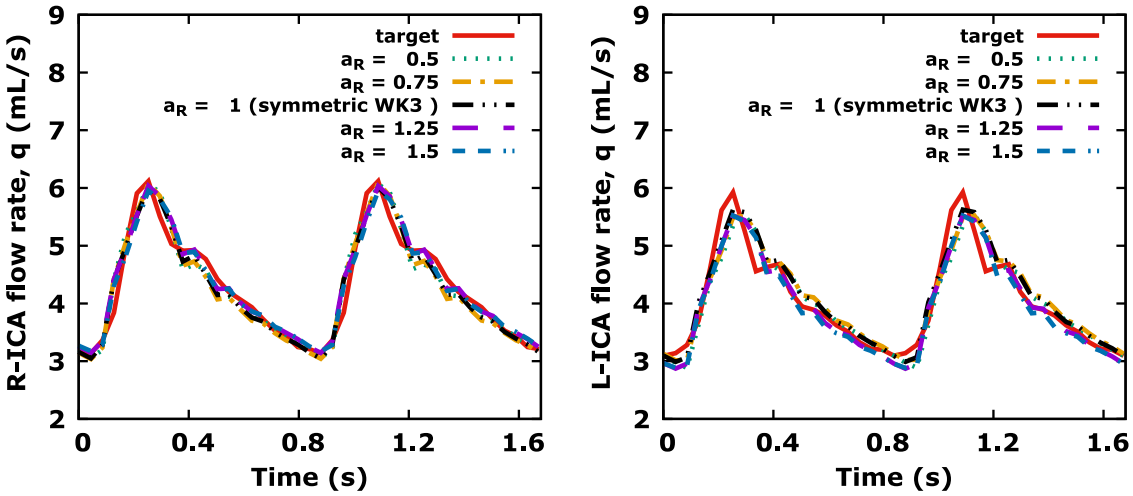

Figure 4.16: Flow rate waveform using WK3 with the same compliances and different resistances for left and right pairs of terminal compartments and comparison with the target flow rate in ICAs.

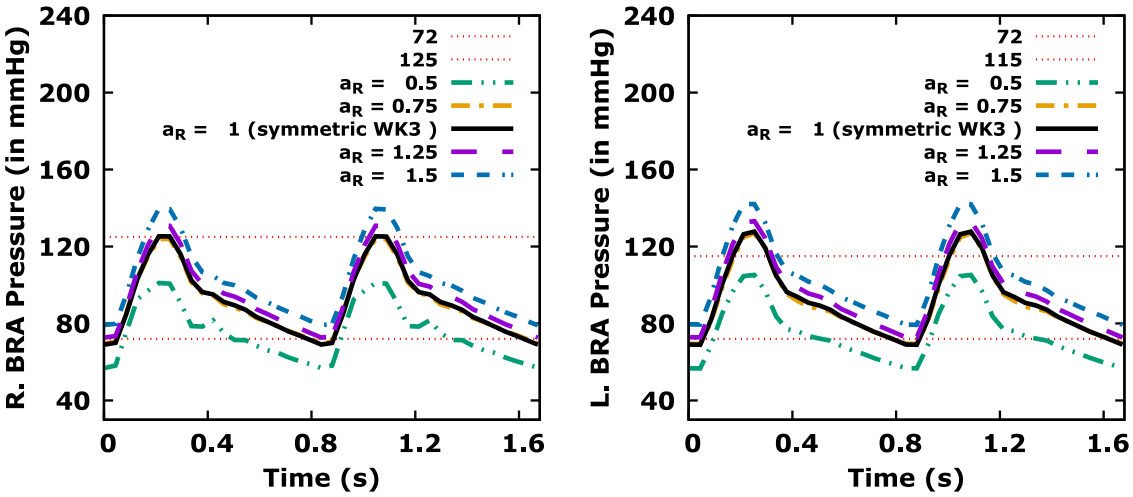

Figure 4.17: Comparison between R. and L. brachial pressures estimated using WK3 with the same compliances and different resistances for left and right pairs of ICAs' terminal compartments. The clinically measured SBP and DBP in the right and the left brachial artery are 125 and 72, and 115 and 72, respectively.

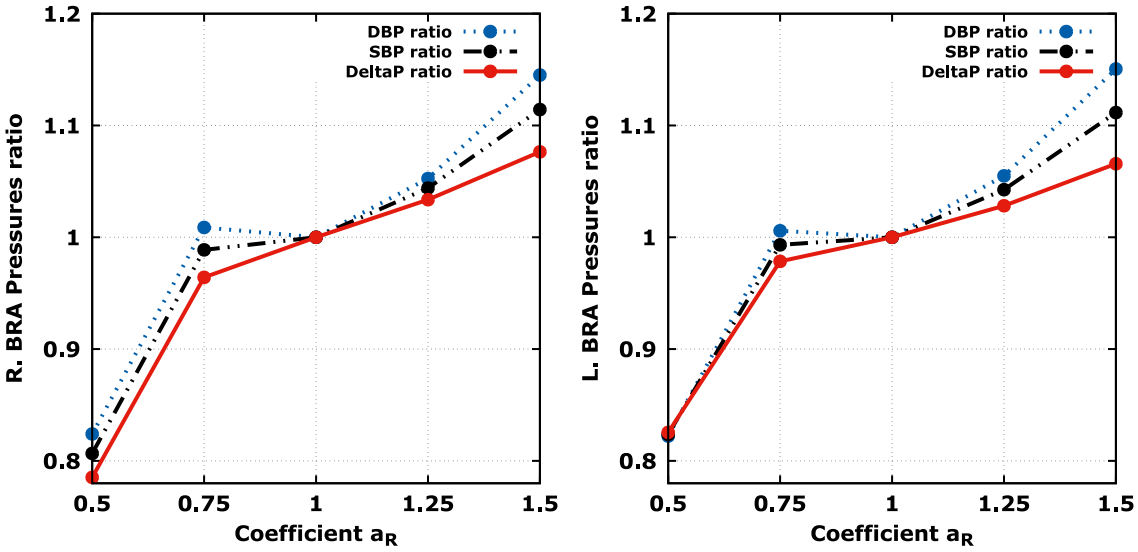

Figure 4.18: Rate of variation of different pressure information with respect to a_R. The different pressure information increase with the level of resistances.

appears that, unlike with the nonsymmetry in the resistances, a loss of symmetry on the compliances has much less impact on the inversions.

As a conclusion, nonsymmetry in the resistances seems to play a much bigger role.

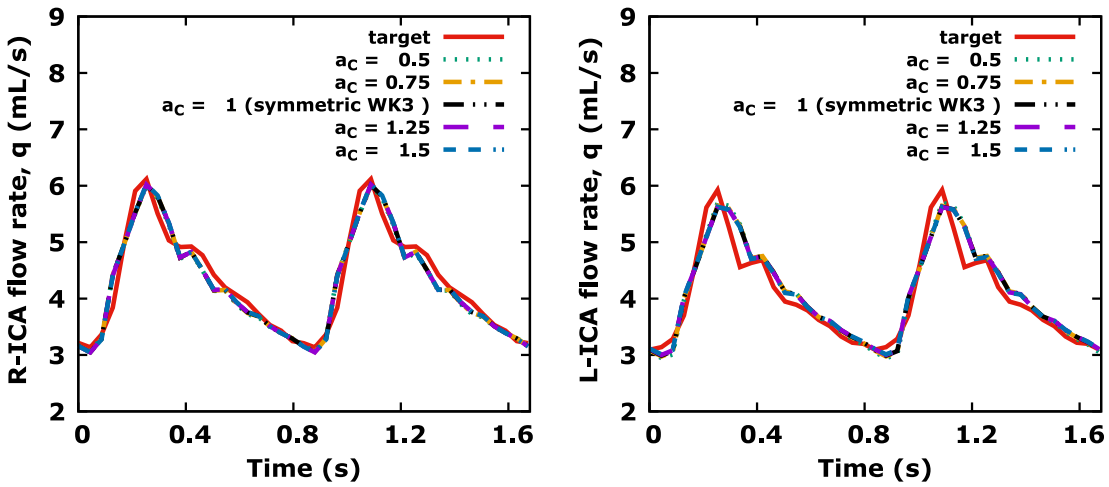

Figure 4.19: Comparison of the model simulated blood flow rate waveform using WK3 with the same resistances and different compliances for left and right pairs of ICAs' terminal compartments.

Finally, both the resistances and compliances are considered as nonsymmetric. Figures 4.22, 4.23, and 4.24 show the impact of this loss of symmetry. One sees that

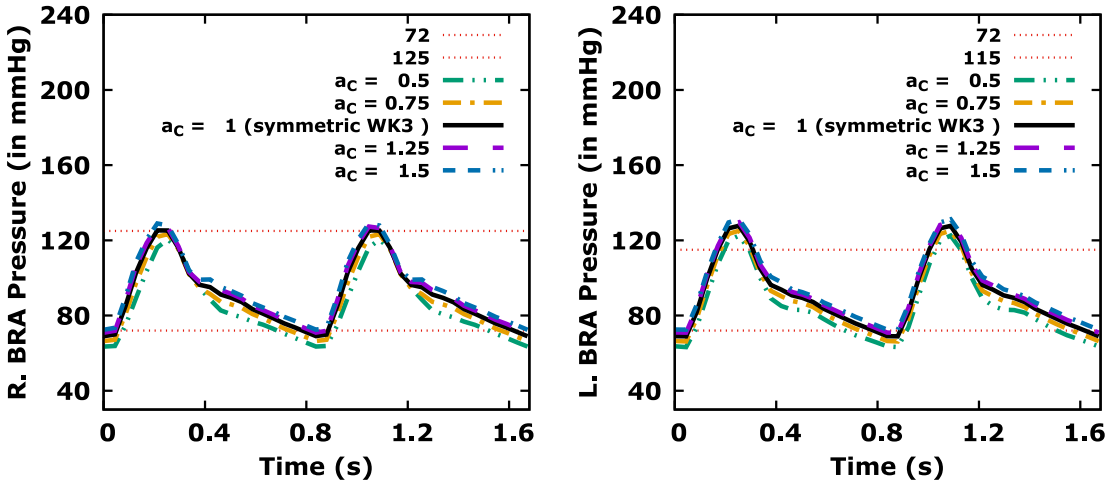

Figure 4.20: Comparison between R. and L. brachial pressures estimated using WK3 with the same resistances and different compliances for left and right pairs of ICAs' terminal compartments. The clinically measured SBP and DBP in the right and the left brachial artery are 125 and 72, and 115 and 72, respectively.

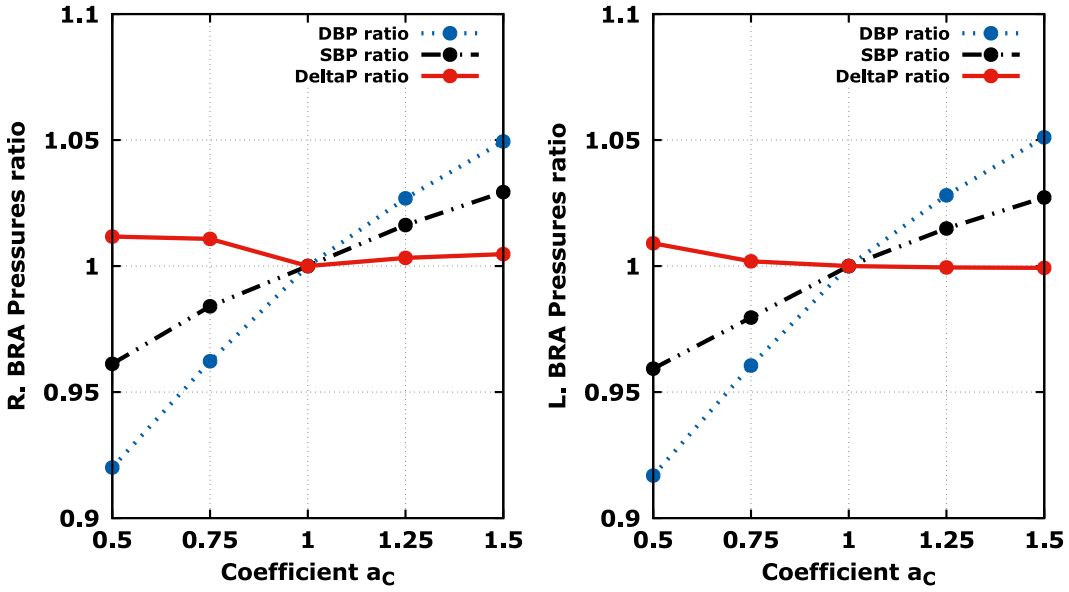

Figure 4.21: Rate of variation of different pressure information with respect to a_C. Despite the systolic and diastolic pressures vary, their difference `DeltaP` is nearly invariant with respect to this nonsymmetry.

the target signals are well recovered by all the inversions. It is therefore unclear if nonsymmetry should be introduced during the inversion if the only available information is the flow rate. Considering the brachial pressures, still the improvement is not spectacular compared to the symmetric configuration as many combinations produce acceptable brachial pressure levels. A symmetric set up appears to be suitable for this patient. This is shown in Figure 4.24 indicating the acceptable nonsymmetry combinations for (a_R, a_C), i.e., the combinations for which similar results than the symmetric WK3 conditions are obtained. The same trends were observed for the other patients considered in this study. In short, if one wants the inversion to identify possible nonsymmetry, one would need more discriminating data.

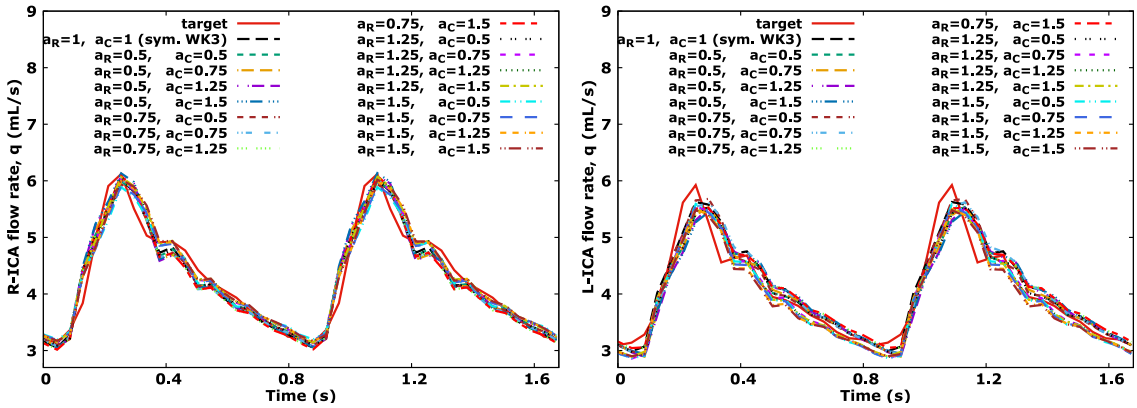

Figure 4.22: Comparison of the model simulated blood flow rate waveform after inversion with different combinations of a_C and a_R.

4.6 Alternative backward uncertainty quantification

4.6.1 Linear theory for an alternative covariance matrix construction

EnKF is an elegant way for backward uncertainty propagation as one has access to the covariance matrix ($\mathbf{cov_x}$) of the optimization variables during the inversion from ensemble information. Indeed, at each iteration k of the EnKF algorithm, $\mathbf{cov_{x_k}}$ can be defined with a similar expression to those used for calculating the error covariance

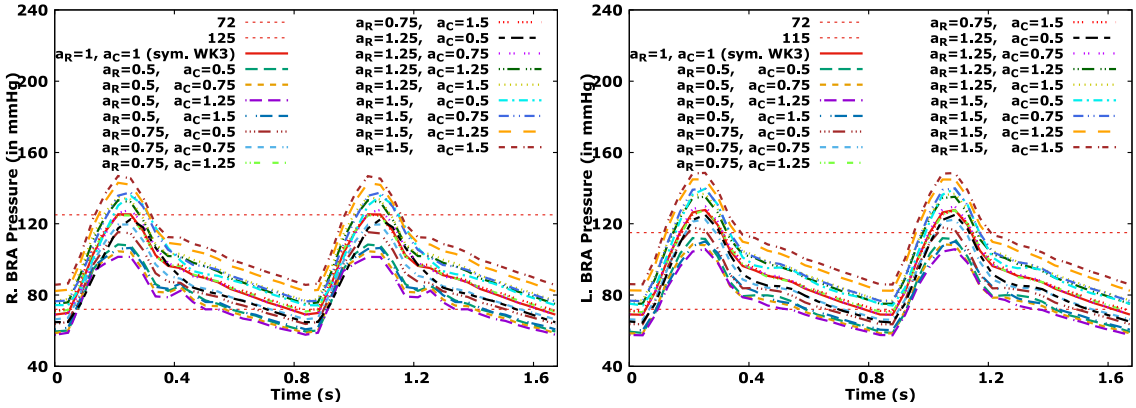

Figure 4.23: Comparison between R. and L. brachial pressures after inversion with different combinations of a_C and a_R.

matrices necessary to the Kalman gain matrix computation:

$$\mathbf{cov}_{\mathbf{x_k}} = \frac{1}{q_{\text{ens}} - 1} \sum_{i=1}^{q_{\text{ens}}} \left[\mathbf{x}_k^{\text{f}_i} - \overline{\mathbf{x}}_k^{\text{f}} \right] \left[\mathbf{x}_k^{\text{f}_i} - \overline{\mathbf{x}}_k^{\text{f}} \right]^\top,$$

where q_{ens} is the size of the ensemble, the superscripts f and f_i stand for forecast and the ith forecast member of an ensemble of size q_{ens}, $\mathbf{x}_k^{\text{f}_i}$ is the forecast ensemble of parameters for $i = 1, \ldots, q_{\text{ens}}$, and $\overline{\mathbf{x}}_k^{\text{f}}$ is the mean of the forecast ensemble of parameters ($\overline{\mathbf{x}}_k^{\text{f}} = \frac{1}{q_{\text{ens}}} \sum_{i=1}^{q_{\text{ens}}} \mathbf{x}_k^{\text{f}_i}$).

This section presents an alternative low-complexity construction of the covariance matrix [325] of the parameters ($\mathbf{cov}_\mathbf{x}$) not necessitating ensemble information. It is based on the assumption of a local linear relationship ($\delta\mathbf{y} = \mathcal{J}\ \delta\mathbf{x}$) between the hemodynamic parameter variations $\delta\mathbf{x} = \mathbf{x} - \overline{\mathbf{x}}$ and the state variations $\delta\mathbf{y} = \mathbf{y} - \overline{\mathbf{y}}$ with respect to the mean values $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$. One interest of this construction is that it can also be used with deterministic minimization algorithms suitable for large dimensional problems where ensemble methods might fail.

Let us establish the expression for the covariance matrix $\mathbf{cov}_\mathbf{x}$ of the parameters $\mathbf{x}$ knowing the covariance matrix $\mathbf{cov}_\mathbf{y}$ of the model solution. These are supposed independent which means that $\mathbf{cov}_\mathbf{y}$ is diagonal.

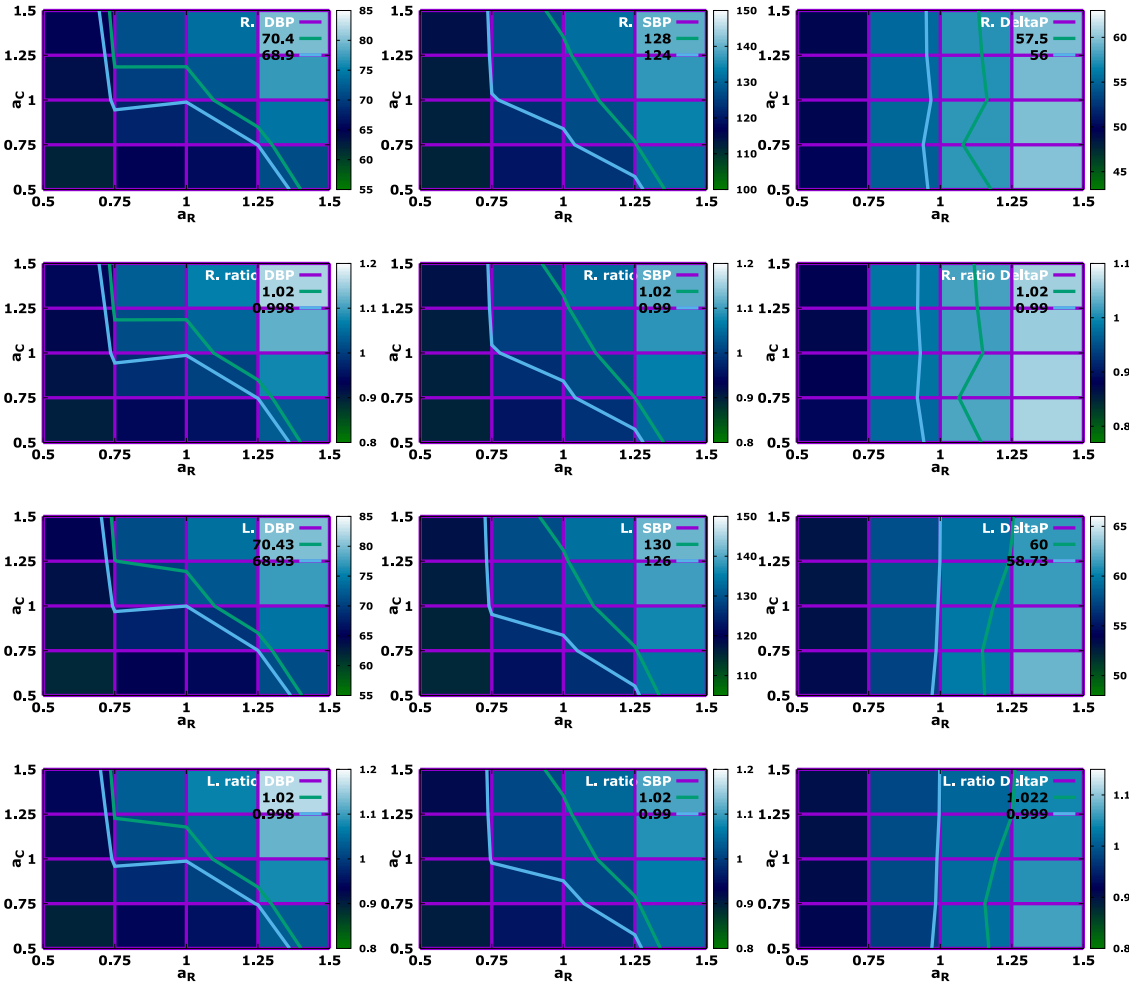

Figure 4.24: Impact of the nonsymmetry on compliances and resistances on the right (R) and left (L) brachial pressures. Diastolic (left column) and Systolic (middle column) pressures (DBP and SBP) are sensitive to nonsymmetry in both resistance and compliance while `DeltaP`, the difference between the two (right column) only sees the nonsymmetry in the resistance. The delimited region by the two contours indicates where the nonsymmetric combination produces similar results than the symmetric WK3 conditions.

The covariance matrix of $\mathbf{y}$ is given by

$$\mathbf{cov_y} = \mathbb{E}\left[(\mathbf{y}-\overline{\mathbf{y}})(\mathbf{y}-\overline{\mathbf{y}})^\top\right] = \mathbb{E}\left[\mathcal{J}\ (\mathbf{x}-\overline{\mathbf{x}})\ (\mathbf{x}-\overline{\mathbf{x}})^\top\ \mathcal{J}^\top\right] = \mathcal{J}\ \ \mathbb{E}\left[(\mathbf{x}-\overline{\mathbf{x}})(\mathbf{x}-\overline{\mathbf{x}})^\top\right]\ \ \mathcal{J}^\top$$
$$= \mathcal{J}\ \ \mathbf{cov_x}\ \ \mathcal{J}^\top,$$

and therefore,

$$\mathbf{cov_x} = \mathcal{J}^{-1}\ \ \mathbf{cov_y}\ \ \mathcal{J}^{-\top},$$

where $\mathcal{J}^{-\top} = (\mathcal{J}^\top)^{-1}$ and $\mathcal{J} = \nabla_{\mathbf{x}}\mathbf{y}$. As $\mathbf{x}$ and $\mathbf{y}$ do not have the same size, $\mathcal{J}$ is a rectangular matrix. A least square formulation is introduced to give a sense to $\mathcal{J}^{-1}$ through its normal form, minimizing

$$\frac{1}{2} < \mathcal{J}\ \ \mathbf{cov_x}\ \ \mathcal{J}^\top, \mathcal{J}\ \ \mathbf{cov_x}\ \ \mathcal{J}^\top > - < \mathbf{cov_y}, \mathcal{J}\ \ \mathbf{cov_x}\ \ \mathcal{J}^\top > .$$

First order optimality condition with respect to $\mathbf{cov_x}$ gives

$$\mathcal{J}^\top\mathcal{J}\ \ \mathbf{cov_x}\ \ \mathcal{J}^\top\mathcal{J} - \mathcal{J}^\top\ \ \mathbf{cov_y}\ \ \mathcal{J} = 0,$$

which implies

$$\mathbf{cov_x} = (\mathcal{J}^\top\mathcal{J})^{-1}\ \ \mathcal{J}^\top\ \ \mathbf{cov_y}\ \ \mathcal{J}\ \ (\mathcal{J}^\top\mathcal{J})^{-1},$$

and finally leads to

$$\mathbf{cov_x} = \mathcal{J}^{-1}\ \ \mathbf{cov_y}\ \ \mathcal{J}^{-\top} = \left(\mathcal{J}^\top\ \ \mathbf{cov_y}^{-1}\ \ \mathcal{J}\right)^{-1}. \tag{4.7}$$

To get $\mathbf{cov_x}$ knowing $\mathbf{cov_y}$, it is therefore sufficient to evaluate the Jacobian $\mathcal{J}$. This can be done using a finite difference approximation.

When the inversion is successful, the model states $\mathbf{y}$ and the observations $\mathbf{y_{obs}}$ are close. Assuming that $\mathbf{cov_{y_{obs}}} \sim \mathbf{cov_y}$, this construction from $\mathbf{cov_y}$ can be extended to $\mathbf{cov_{y_{obs}}}$.

4.6.2 Application to the patient-specific data

This section applies the backward uncertainty quantification construction presented in Section 4.6.1 to the patient-specific data and compares the outcome with the ensemble analysis through EnKF results.

Figure 4.25 shows a comparative overview of the square root of $\mathbf{cov_x}$ diagonal elements calculated using EnKF (blue points) and the low cost linear theory (red points). These values represent standard deviation associated with the estimated parameters. The linear theory seems to provide lower bounds for the backward uncertainty on the inversion. This comparison shows that the linear hypothesis permits to *a posteriori* obtain the uncertainty on the inversion solution with an error of less than 15%, and this especially without requiring any extra information when used together with a gradient based minimization algorithm.

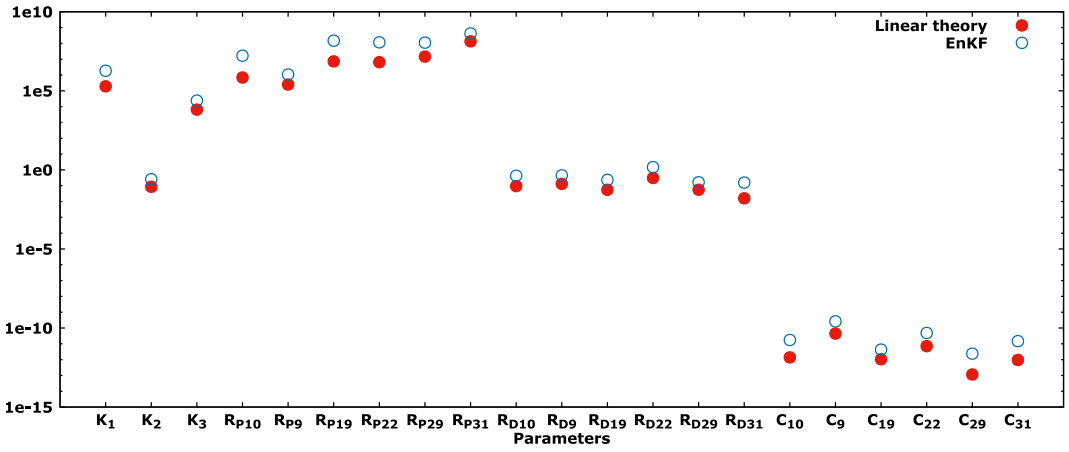

Figure 4.25: Comparison of the square root of $\mathbf{cov_x}$ diagonal elements coming from EnKF solution (blue points) to the square root of $\mathbf{cov_x}$ diagonal elements calculated with the low complexity deterministic evaluation based on the linear theory given in equation 4.7 (red points).

4.7 Limitations

It is worth bearing in mind the limitations of the present study, both intentional and unintentional. One of the limitations could be the use of 0D models. Indeed, as pointed out by Shi et al. [91] and Kokalari et al. [92], the 0D models consider as uniform the distribution of fundamental variables (pressure, volume, and flow rate) at any instant in time. Thereby, they do not account for the spatial variation of these parameters [92]. They are therefore only appropriate for the study of global distribution of the pressure,

flow rate, and blood volume in systemic arterial network and for specific physiological conditions. Phenomena, such as wave transmission, wave travel, and wave reflections, cannot be or are not studied [326]. It is worth noting that several studies have pointed out the shortcomings of the lumped models and have proposed more complex and sophisticated models to better capture the characteristics of blood flow [91, 316, 327, 328, 329]. However, this is not central to our discussion and is beyond the scope of this paper. What is shown here and is regularly employed in the literature concerning arterial hemodynamics is that this simple model permits to give reasonable results. Moreover, the region studied being compact, (the size of the cerebral network [of order 10 cm] is small compared with typical pressure wavelength [of order a few meters]), assuming that all the quantities are homogeneous over each segment (viz, making the 0D assumption at the scale of each element), is acceptable. Of course, it would not be acceptable if this study was dealing with the arterial network of the whole body.

Another limitation could be the choice of the formula relating Young modulus, thickness, and radius for the cerebral arteries in Equation 4.4. This work assumes the same phenomenological relation for both large and cerebral arteries. The other factor might be the failure to account for the variation of arteries radii along arterial axes and the adaptation of missing data, such as the geometrical parameters of the arteries, from the literature. Future work should include sensitivity analysis of the optimization parameters outcomes with respect to this simplification. For instance, one way to quantify how this simplification would affect the results of optimization parameters is to perform several inversions, with the present tool, using perturbed geometries and see the impact of this perturbation on the outcomes. This is a step to be taken in the future to improve the robustness of the approach.

As possible improvements, one could also mention the marked changes in simulated brachial pressure wave resulting from assigning different resistances for left and right pairs of terminal cerebral compartments. From the physiological point of view, this might not be correct as some studies support that arterial blood pressure wave is

mainly determined by the systemic vessels rather than the cerebral vessels. In other words, changing cerebral vascular resistances should not induce such large changes in systemic arterial pressure wave. To demonstrate the relevance of this issue, new tests of nonsymmetry in resistances with $a_R = 1.5$ only have been performed. The results are summarized in Figure 4.26. They show that changing the resistances of only one termination in cerebral or brachial arteries has no effect on brachial pressure. However, they reveal that changing resistances at only thoracic aorta termination has the same impact (or even more) on brachial pressure as the nonsymmetry in resistances at all the arterial terminations except thoracic aorta case above presented. Furthermore, one notices that changing resistances at all terminations induces more important changes in brachial pressure. Thus, the apparent effects of resistances at arterial terminations on the brachial pressure might result from the fact that the lower part of systemic circulation, which is of major importance in determining arterial blood pressure wave is only represented here by distal boundary conditions.

Also, concerning the sensitivity to change in the outlet boundary condition, one would expect the mean values to be achievable through WK1 (but not necessarily the pulse pressure). Indeed, as pointed out by Vignon and Taylor [330], a resistance boundary condition damps the flow waveform and increases the pulse pressure but does not change the mean values.

At last, it is necessary to highlight the lack of clinically measured flow rates in arteries other than ICAs and ascending aorta. This prevents the proper validation of the predicted flow divisions among major tissues. More clinical data are required.

4.8 Conclusion

The robustness of a functional cerebral imaging platform has been analyzed for patient-specific situations. In particular, detailed results for one patient have been reported to illustrate the kind of information and sensitivity analyses which can be produced in addition to the classical inversion results. The aim is to be able to quantify the impact

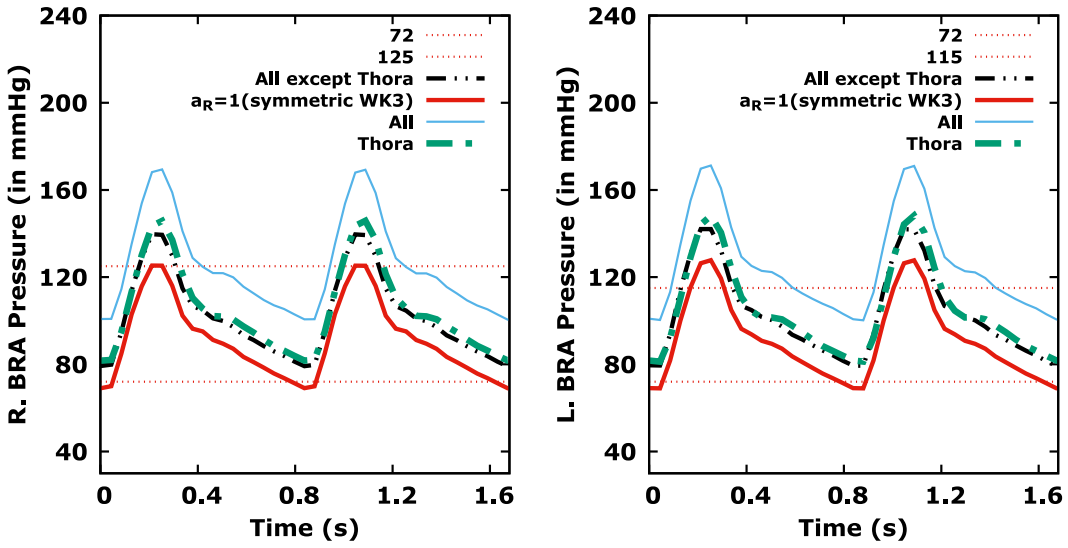

Figure 4.26: Importance of systemic arteries with respect to cerebral arteries in determining R. and L. brachial pressures estimated using WK3 with nonsymmetry in resistances with $a_R = 1.5$ only. Changing resistances at only thoracic aorta termination (dashed green) has the same impact on brachial pressure as the nonsymmetry in resistances at all the arterial terminations except thoracic aorta case (dashed black). Changing resistances at all terminations (thin blue) induces more important changes in brachial pressure. The clinically measured SBP and DBP in the right and the left brachial artery are 125 and 72, and 115 and 72, respectively.

of the uncertainty of different parameters usually considered as deterministic on the hemodynamic parameters obtained after inversion.

An important source of uncertainty being the flow rate from the heart is that the sensitivity of the inversion with respect of the inlet flow has been analyzed. The conclusion drawn from this analysis has been that the predicted pressure and the elastic parameter estimation are of the same order of magnitude than the assumed uncertainty on the inlet flow rate. This provides additional guidance to practitioners as it gives some indication for the level of uncertainty one can tolerate because of manual acquisition steps. The sensitivity of the inversion with respect to the choice of the boundary conditions in the direct model has been analyzed. The WK3 boundary conditions appear necessary for a correct retrieval of the target signals. Then the impact of nonsymmetry in the terminations resistances and compliances has been analyzed. The former effects appear prominent. However, one notices that available

observations are not enough discriminating and, therefore, do not make it obvious for possible nonsymmetry to be identified through the inversion procedure.

Finally, an ensemble approach might be unsuccessful if the size of the inversion problem is large. To address this issue, one might use deterministic gradient-based inversion algorithms together with an adjoint formulation for the gradient evaluation. But these algorithms do not propagate backward the uncertainty on the observations. To address this issue, a low-complexity backward uncertainty quantification construction has been introduced. It uses a local linear relationship between the blood flow rate and the hemodynamic parameters. The construction has been favorably compared to the ensemble outcome, which is interesting as this is a generic construction and can be used with any deterministic inversion algorithm.

Acknowledgment

The authors greatly thank Prof-Dr V. Costalat, Dr E. Le Bars and Dr J. Deverdun from the Department of Neuroradiology of the *Centre Hospitalier Régional Universitaire de Montpellier, Gui de Chauliac, Montpellier, France*, for providing the medical data and Dr J. Siguenza and Dr R. Lal for helpful discussions.

Chapter key points:

- **Estimated blood pressure and elastic parameters are of the same order of magnitude than the assumed uncertainty on the inlet flow rate.**
- **The Windkessel WK3 boundary conditions appear necessary for a correct retrieval of the target signals.**
- **Available observations are not enough discriminating to the extent that a possible nonsymmetry in the boundary condition parameters can be identified through the inversion procedure.**
- **A low-complexity backward uncertainty quantification construction is introduced and favorably compared with the ensemble outcomes.**

CHAPTER 5

Global sensitivity analysis for assessing the parameters importance and setting a stopping criterion in a biomedical inverse problem

Chapter contents

This chapter has been submitted for publication as:

Rapadamnaba R., Ribatet M., Mohammadi B. "Global sensitivity analysis for assessing the parameters importance and setting a stopping criterion in a biomedical inverse problem" in *International Journal for Numerical Methods in Biomedical Engineering*, 2020 .

The numbering of sections, figures and tables in the original manuscript have been altered for this chapter, to be consistent with the thesis chapter numbering. The references have been incorporated into the global references at the end of the thesis.

Part of this chapter has been presented in the previous chapter as follows:

- Section 5.2 presents a summary of Sections 4.2 and 4.3 of chapter 4,
- Figures of Section 5.3 are taken from Section 4.5.1 of chapter 4.

The rest of the sections of this chapter presents new materials focusing on several sensitivity analyses of the EnKF-based inversion procedure. The aim of this chapter is to rank blood flow model parameters previously estimated using the inversion procedure by order of importance and show how to set a stopping criterion for the EnKF algorithm used in the inversion. A linear model is assumed in order to easily perform sensitivity analysis using the Sobol' indices technique. An interesting *a posteriori* consequence of this analysis is that this simplified model permits to understand that most of proximal resistances are relevant parameters for blood flow rates estimation in internal carotid arteries.

Abstract

This paper shows how to obtain in addition to the standard deviations available after an ensemble Kalman filter based inversion procedure, an apportioning of the total uncertainty in the outputs of a patient-specific blood flow model into small portions of uncertainty due to the model input parameters. Some indicators often used for analyzing the behavior of numerical parameters, such as the variability of parameters over time, the standard deviations, the coefficients of variation, and the Sobol' first order and total indices are discussed. These allow the identification of the importance rank of the different input parameters both for an analytical model and for the patient-specific blood flow model, as well as the influence of the interactions between these parameters on the output variance of both models. The results obtained using the patient-specific blood flow model show that knowing the importance rank of the model input parameters is useful for setting a suitable stopping criterion for the algorithm used in the inversion procedure in order to avoid unnecessary over-solving, mainly in clinical situations where faster diagnosis is always requested. Indeed, the work permits to reduce typically by a factor of six the time to solution and most importantly with very limited extra calculation using already available information.

Keywords: sensitivity analysis, Sobol' sensitivity indices, uncertainty quantification, hemodynamic inverse problems, parameter estimation, convergence stopping criterion.

5.1 Introduction

Models have always been invaluable tools used to explain a phenomenon or a system, to study the effects of their different components, to communicate knowledge, and to make decision or predictions about the future behavior of a real-world phenomenon [331]. In short, models describe our beliefs about how the world functions. Of all models, mathematical models, which translate those beliefs into

mathematical language — namely the description of a system by a set of variables and a set of equations establishing relationships between the variables — are of great importance in the research world. Indeed, mathematical modeling, or in the words of Neumaier, "the art of translating problems from an application area into tractable mathematical formulations whose theoretical and numerical analysis provides insight, answers, and guidance useful for the originating application" [332], is inseparable from most of scientific research disciplines to some extent. In fact, most often providing a simplified version of something that is real, mathematical models are indispensable in many fields of application [332, 333]. For instance, in the field of applied mathematics, applied physics, and in many other fields, mathematical modeling is the starting point for analyzing a system to be understood, controlled or optimized. This allows researchers to build descriptive models of the system as a hypothesis of how the system could work, or try to estimate how an unexpected event could affect the system, and even try out different control approaches of the system using numerical simulations.

Unfortunately, a mathematical model can have or includes significant spurious artifacts, such as random events, measurement errors, more-or-less uncertain parameters, missing information, unrealistic nonlinearities, and also conceptual uncertainty, i.e., poor knowledge or partial, erroneous understanding of the driving forces and mechanisms in the model, or uncertainty in the model structures, assumptions and specifications [162, 333, 334, 335]. Moreover, it can particularly be highly complex, and as a result, the relationships between its inputs and its outputs may be poorly understood. In this context, as pointed out by Saltelli, issues such as relevance and transparency become critical [166]. Consequently, good modeling practice requires that the modeler provides an evaluation of the confidence in the model [333]. This requires, on the one hand, uncertainty analysis in the model results, and on the other, an evaluation of the part of output uncertainty due to each model input or subgroup of the model inputs.

This work only focuses on the latter issue, which is the prerogative of sensitivity analysis (SA). After quantifying uncertainties in a 0D arterial blood flow model of patient-specific upper body region including the circle of Willis in order to test its robustness and its relevance with respect to some choices and assumptions made in the model in previous works [12, 153, 277], this paper answers the question, which input variables or which input combinations are causing the largest effect on the model output? In other words, this work aims at addressing the following issues [336]: which input parameters cause the largest variation in the output? Is there any input parameter whose variability has a negligible effect on the output? Are there interactions that amplify or dampen the variability induced by individual parameters? Specifically and in the words of Saltelli, this study shows how the uncertainty in the model output can be apportioned to different sources of uncertainty in the model inputs [166]. This type of approach, also known as importance measures or sensitivity indices, is known to be a variance-based method, which provides a factor-based decomposition of the output variance and implicitly assumes that this moment is sufficient to describe the output variability [166]. Therefore, the main objective of our approach is to identify and rank in order of importance the important factors in the blood flow model in order to establish a stopping criterion for the ensemble Kalman filter (EnKF) based inversion algorithm used in previous works [12, 277]. To be more precise, this work employs some indicators commonly used for inspecting the behavior of numerical parameters, such as the variability of parameters over time, the standard deviations, the coefficients of variation, and the Sobol' first order and total indices to analyze the link between the convergence of the 0D arterial blood flow model parameters estimated in previous works [12, 277] using an iterative inversion procedure and the importance of these parameters through SA with the following steps:

- Proceed with the inversion until the model with identified parameters fit the target observations,

- Identify the variables in the model which do not fully converge although target

observations have been fitted in order to rank them by order of importance using alternative sensitivity evaluations,

- Find a stopping criterion for the inversion in order to perform faster diagnosis in clinical cases.

The importance and usefulness of SA is no longer debatable. It is widely recognized that the knowledge of sensitive input parameters can be beneficial for a range of purposes [334], such as better understanding of the relationships between input and output variables in a model, testing the robustness of a model results, development of robust and better model by reducing errors and uncertainties in the model, or by simplifying the model [335, 337, 338, 339]. Many publications on the subject, in particular the works of Saltelli et al. and references therein [160, 166, 340, 341, 342], explain and illustrate these objectives.

The layout of the paper is as follows. The biomedical problem addressed and the blood flow model considered in this study are reported in Section 5.2. Section 5.3 presents in detail the motivations of the study while Section 5.4 provides some basic sensitivity evaluations based on the blood flow model results. Section 5.5 introduces the state of the art of sensitivity analysis with particular emphasis on the Sobol' indices for independent input variables, as well as on their estimation method. Section 5.6 presents two applications illustrating the usefulness and importance of the sensitivity indices, first in an analytical model and then in a patient-specific blood flow model, while Section 5.7 is intended to provide concluding remarks.

5.2 Materials and problem definition

5.2.1 Arterial network structure and model equations

Figure 5.1 shows the 0D arterial model (Figure 5.1B) used in this study together with its equivalent 1D model (Figure 5.1A) adapted from Alastruey [284]. It represents the arteries of a patient-specific upper body, including those of the circle of Willis (coW).

Each segment in this arterial network structure and in the 0D models in general is analogous to an electrical circuit [121, 305, 315, 316], and is therefore represented as shown in Figure 5.2. For further details, the reader is referred to previous studies [12, 277, 278].

Thus, applying the Kirchhoff's current and voltage laws to a single arterial segment assumed being filled with an incompressible Newtonian fluid, the following system of coupled ordinary differential equations for blood pressure and blood flow rates is obtained [121, 315, 316] :

$$\begin{array}{rcll} \dfrac{dP_{out}}{dt} & = & \dfrac{q_{in} - q_{out}}{C} & \text{(Pressure equation)} \\ \dfrac{dq_{in}}{dt} & = & \dfrac{P_{in} - P_{out} - Rq_{in}}{L} & \text{(Flow equation),} \end{array} \tag{5.1}$$

where for each arterial segment, P_{in}, P_{out}, q_{in}, q_{out}, R, L, and C are inlet blood pressure, outlet blood pressure, inlet flow rate, outlet flow rate, the viscous flow resistance, blood inertia, and the artery compliance, respectively. For arteries with a radius $r < 0.2$ cm, the inertial effect is ignored [317], and in this case, the flow equation is given by $q_{in} = (P_{in} - P_{out})/R$.

The electrical parameters (R, L, and C) of each arterial segment are related to the geometrical and physical parameters (r, h, l, and E) as follows [305] :

$$R = \frac{8\mu l}{\pi r^4} \text{ (Hagen-Poiseuille equation)}, \quad L = \frac{\rho l}{\pi r^2}, \quad \text{and} \quad C = \frac{3\pi r^3 l}{2Eh},$$

where r, h, l, μ, ρ, and E are the radius of the artery, the arterial wall thickness, the length of the arterial segment, the blood viscosity, the blood density, and the Young modulus, respectively.

At the bifurcations and at the junctions of the arterial trees shown in Figure 5.1, the enforcement of the mass and momentum conservation laws makes it possible to prescribe for diverging and merging flows the conditions at the nodes, namely that the blood pressure remains continuous and the blood flow rates must be conserved. For further details, the reader is referred to [126, 322, 343].

In order to take into account the effect of the downstream vasculature, the three-element Windkessel model (WK3-lumped parameter model) [286, 318] is applied at the outlet of each terminal arterial segment. As a reminder, the governing ordinary differential equation relating the instantaneous blood pressure and the blood flow rate for the WK3 model is given as

$$\frac{dp(t)}{dt} + \frac{p}{R_D C} = R_P \frac{dq(t)}{dt} + \frac{qR_T}{R_D C}, \tag{5.2}$$

where p, q, R_P, R_D, and C, are the instantaneous pressure at the inlet of the WK3 model, the instantaneous flow rate, the proximal resistance, the distal resistance, and the compliance of the vascular bed, respectively.

To solve the arterial network equations (the system of first-order ordinary differential equations consisting of Equations (5.1) and (5.2), the Fortran version of an implicit numerical integration solver `DVODE` [319, 320], available on `http://www.radford.edu/~thompson/vodef90web/`, is used.

5.2.2 Problem definition

Given the physical model described above, this study focuses on the following optimization problem:

$$\min_{\mathbf{x}\in\mathbf{O}_{ad}} J\big(t, \mathbf{y}(\mathbf{x},\mathbf{z},\mathbf{t}), \mathbf{y_{obs}}(\mathbf{t})\big), \tag{5.3}$$

$$\mathbf{y}(\mathbf{t}) \ \text{ and } \ \mathbf{y_{obs}}(\mathbf{t}) \in \mathbb{R}^2,\ \mathbf{z} \in \mathbb{R}^3,\ \mathbf{x} \in \mathbf{O}_{ad} \subset \mathbb{R}^{21},$$

where $\mathbf{x}$, $\mathbf{y_{obs}}(\mathbf{t})$, and $\mathbf{z}$ are independent variables and the time-dependent cost function J, of the form:

$$J\big(t, \mathbf{y}(\mathbf{x},\mathbf{z},t), \mathbf{y_{obs}}(t)\big) = \frac{1}{2}\|\mathbf{y}(\mathbf{x},\mathbf{z},t) - \mathbf{y_{obs}}(t)\|^2.$$

Only the state variable, $\mathbf{y}(t)$, solution of a state equation $F(\mathbf{y}(\mathbf{x},\mathbf{z},\mathbf{t})) = 0$ and the time variable, t, are related to the cost function, J. The optimization parameter $\mathbf{x}$, consisting of 21 components, which will be specified later in Section 5.6.2.1, belongs to $\mathbf{O}_{ad}$, the optimization admissible domain [311], and represents the hemodynamic parameters for

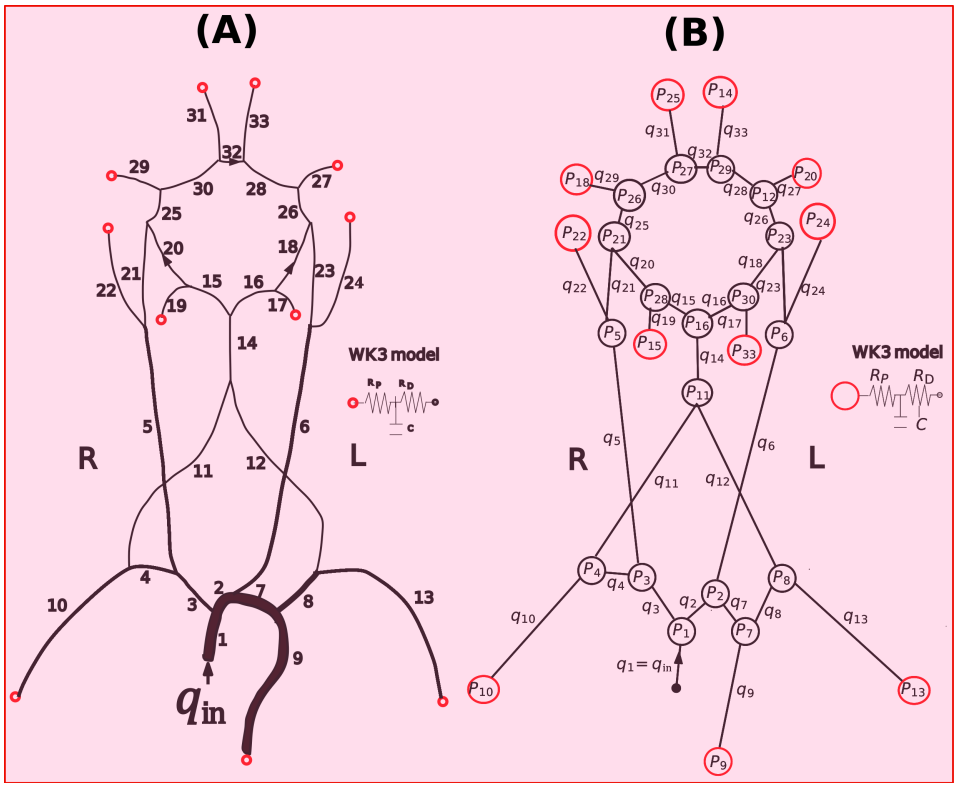

Figure 5.1: The network [284] of a 1D blood flow model of the upper body arteries and of the coW (A), and its equivalent compartment model (B). The lines size and the arrows indicate the relative size of the arteries and the direction of flow, respectively. The eleven red circles in each Figure represent Windkessel WK3 model boundary conditions.

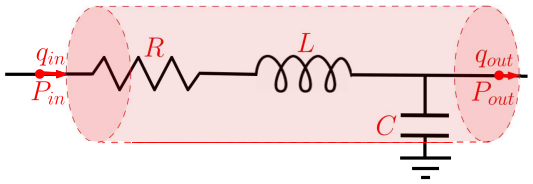

Figure 5.2: A single compartment circuit illustration.

each segment associated with the WK3 model parameters in Figure 5.1. The variable $\mathbf{z}$ contains the 3 parameters used in Rapadamnaba et al. [277] to model the inlet flow rate signal, q_{in}, in Figure 5.1. The observations variable, $\mathbf{y_{obs}}(t)$, represents patient-specific blood flow rates time series obtained from magnetic resonance angiography and magnetic resonance imaging (MRA & MRI) acquisitions at right and left internal carotid arteries (R-ICA & L-ICA) as described in previous studies [12, 277]. The state variable, $\mathbf{y}(t)$, is the estimation of the observation data, $\mathbf{y_{obs}}(t)$, using the blood flow model presented above. The goal of this minimization problem is to find the set of optimization parameters, $\mathbf{x}$, which minimizes the error (the Euclidean distance) between the estimate of the observed data, $\mathbf{y}(t)$, and the observations themselves, $\mathbf{y_{obs}}(t)$. This was the mission of the EnKF-based parameter estimation algorithm used in previous works [12, 153, 277].

5.3 Motivation of the study : typical convergence using the EnKF-based parameter estimation algorithm

The inversion procedure based on the EnKF algorithm used and described in previous works [12, 153, 277] is an elegant way for parameter estimation. In these studies [12, 153, 277], the inversion procedure was coupled either to a 1D arterial network or to a 0D compartment arterial network with the aim of solving the optimization problem recalled above in (5.3). This made it possible, for example, to estimate hemodynamic parameters, which are difficult to identify noninvasively. Thus, in Lal et al. [12, 153], the Young modulus of blood vessels and 21 blood flow model parameters, including the terminal boundary parameters of the arterial network shown in Figure 5.1 were estimated as the solution of an inverse problem, on the one hand, and, on the other hand, blood pressure in cerebral arteries was estimated after inversion. Figure 5.3A,B shows typical evolutions over time of target (observed blood flow rates in R-ICA & L-ICA, $\mathbf{y_{obs}}(t)$) and model (estimated blood flow rates in R-ICA & L-ICA) signals, and Figure 5.3C,D,E,F, typical evolutions of the first two moments for some Young moduli

(Figure 5.3C) and for some WK3 model parameters (Figure 5.3D,E,F), both obtained in previous works [12, 277] using the EnKF-based parameter estimation algorithm.

One observes that although some variables do not fully converge (for example, $R_{D_{22}}$ and C_{22} in Figures 5.3E and 5.3F, respectively), from 8.35 s (i.e., after 10 cardiac cycles), there is a very good agreement between the target and predicted flow rate waveforms meaning that the observations are well recovered (Figure 5.3A,B). It is also noted that iterating more with the inversion algorithm does not modify the recovery of the signals (Figure 5.4C,D), so we would like to stop the inversion procedure earlier to avoid unnecessary calculations. These two findings raise the following interesting issues, which are worth exploring, examining, more closely:

- How to explain the fact that the target signals (observations) are retrieved whereas some estimated parameters have not fully converged?

- Given also that iterating more does not change the signal overlap, when can the inversion procedure be stopped to avoid unnecessary calculations?

One possible response (hypothesis) to the first issue raised above – to be considered as an 'ansatz'– would be that **the parameters not having fully converged, have little or no effect on the cost function, J, (i.e., the error between the estimated and observed blood flow rates), and therefore little or no effect in estimating the observed blood flow rate.** The first way to test this hypothesis is to resort to the variability of estimated parameters over time and also to some basic statistical indicators of dispersion often used in the analysis of a numerical variable, namely standard deviations, coefficients of variation or relative standard deviations, in the hope of getting some clues about the influence of the estimated parameters on the estimate of the observed blood flow rate. The results obtained using these indicators could then be useful in addressing the second issue raised above.

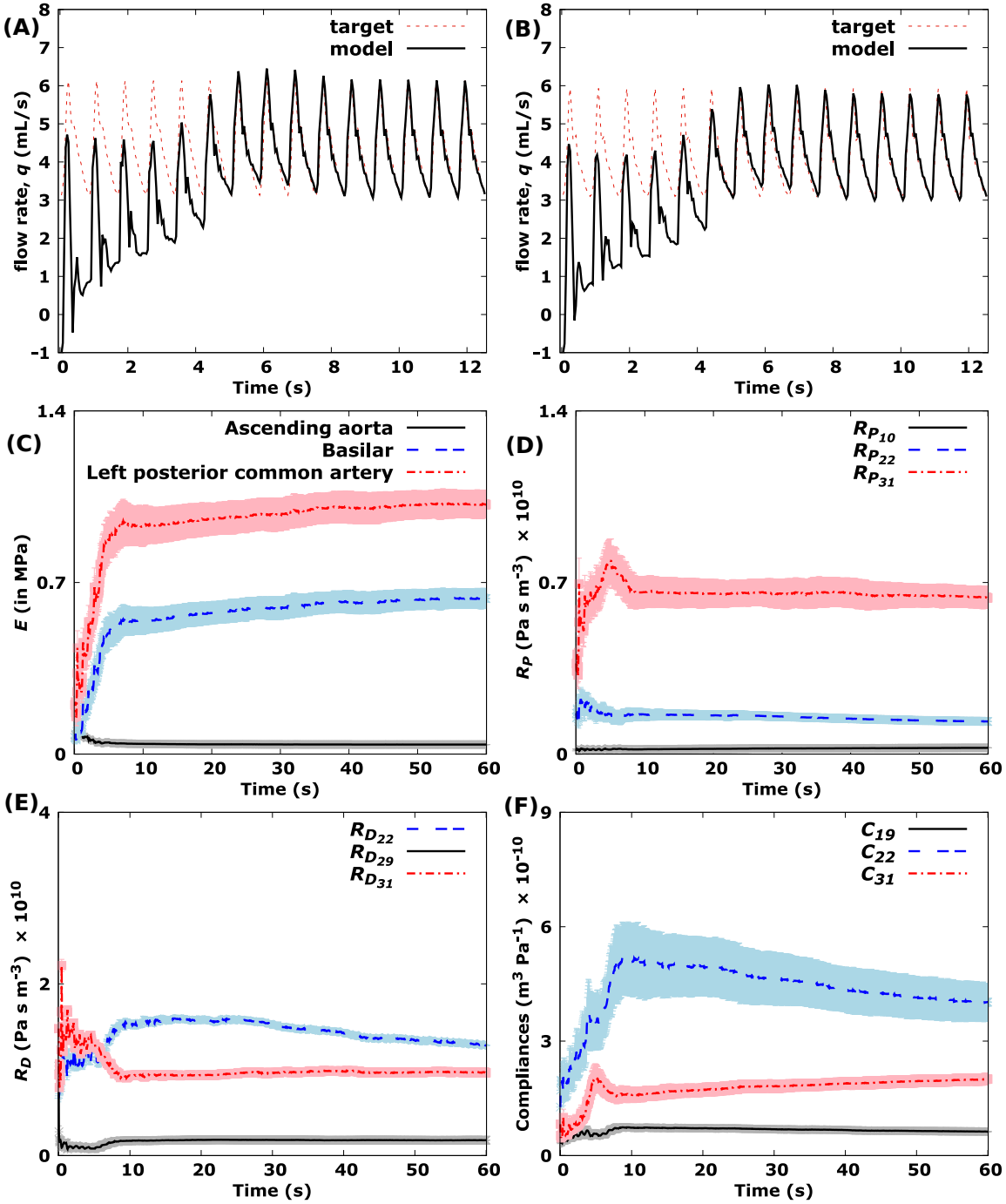

Figure 5.3: Typical convergence of some parameters previously estimated in [277] using the ensemble Kalman filter (EnKF) based algorithm. A,B: Comparison of the model simulated blood flow rate waveform in R-ICA (A) and L-ICA (B) to the target signals. C: Mean and standard deviation of estimated Young moduli for three of the 33 segments of the arterial networks in Figure 5.1. The evolution is over 8.35 s corresponding to about 10 cardiac cycles. D,E,F : Histories of the first two moments of some estimated parameters (Windkessel WK3 model parameters in Figure 5.1) during the EnKF-based algorithm iterations.

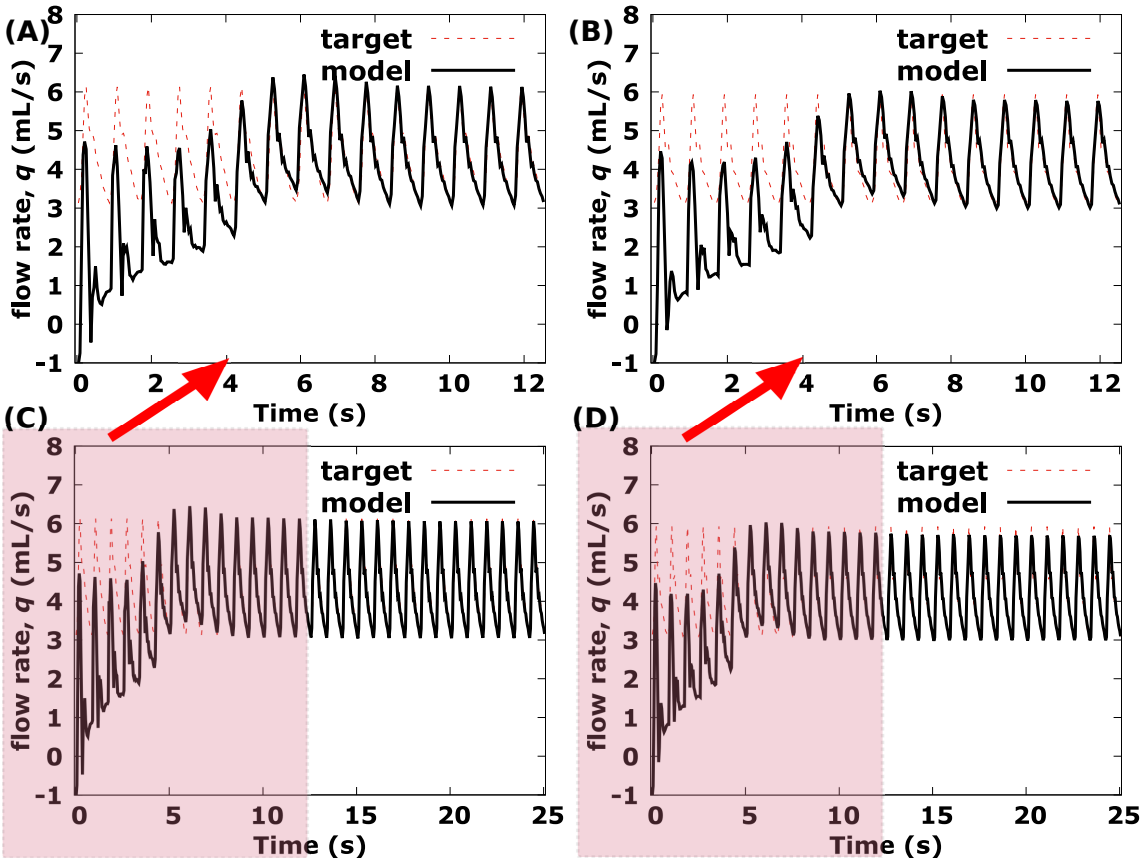

Figure 5.4: Comparison of the model simulated blood flow rate waveform in the left and right internal carotid arteries (L & R ICAs) to the target signals. A,B: comparison of the model simulated blood flow rate waveform in R-ICA (A) and L-ICA (B) to the target signals for the first 12 seconds (approximately 15 cardiac cycles). This is a simple duplication of Figure 5.3A,B and corresponds here to the result of zooming in on the first 12 seconds of C and D. C,D: Comparison of the model simulated blood flow rate waveform in R-ICA (C) and L-ICA (D) to the target signals for the first 25 seconds (approximately 30 cardiac cycles) showing that once the simulated blood flow signal is superimposed on the target one, it does no longer change even with further iterations.

5.4 Hypothesis testing with basic statistical tools

5.4.1 Hypothesis testing using the variability of the optimization parameters during the inversion

To test the hypothesis formulated above, one way is to take a closer look at the convergence of the different estimated parameters by analyzing and comparing the variability of these parameters between two successive EnKF-based algorithm iterations during the inversion procedure. For each parameter i estimated, this variability at a given EnKF-based algorithm iteration k (let us call it V_k^i), is defined as the absolute value of the ratio of the difference between the two values of the i^{th} estimated parameter at the two successive EnKF-based algorithm iterations, k and $k+1$, to the difference between the values of the i^{th} parameter estimated at the first two EnKF-based algorithm iterations. This means that at each iteration k of the EnKF-based algorithm, V_k^i is obtained as follows:

$$V_k^i = \left| \frac{\mathbf{x}_{k+1}^i - \mathbf{x}_k^i}{\mathbf{x}_2^i - \mathbf{x}_1^i} \right|, \; i = 1, \cdots, 21,$$

where i , $\mathbf{x}_k^i$ and $\mathbf{x}_{k+1}^i$ are the estimated parameter number, the i^{th} estimated parameter at the iterations k and $k+1$, respectively. Therefore, $\mathbf{x}_1^i$ and $\mathbf{x}_2^i$ denote the values of the i^{th} estimated parameter at the first two EnKF-based algorithm iterations. This notion of variability is very useful since it provides a much more accurate answer as to whether or not the estimated parameters converge. Indeed, the more unstable the variability of a given estimated parameter is during all the EnKF-based algorithm iterations, (or in other words, the greater the variability of a given estimated parameter varies during the whole inversion procedure), the more the concerned parameter could not fully converge. This means that parameters with very high variability during all the EnKF-based algorithm iterations are those that have not fully converged.

Figure 5.5 shows the variability over time of the 21 parameters estimated in previous works [12, 277]. It can be observed that among the 9 estimated parameters whose the

evolution during the inversion procedure is shown in Figure 5.3D,E,F, only $R_{D_{22}}$ and C_{22} relatively show unstable variabilities during all the EnKF-based algorithm iterations. Hence, $R_{D_{22}}$ and C_{22} are parameters that did not fully converge. This is consistent with the statement made above in Section 5.3 about the non-full convergence of these two parameters but does not say anything about the influence or not of these parameters on the blood flow model output. Thus, the variability simply provides an answer to the question of whether or not the estimated parameters converged, but does not in any way allow us properly to confirm or disprove the initial hypothesis.

5.4.2 Hypothesis testing using standard deviations

Another way to closely evaluate the original hypothesis formulated above could be the use of the well-known basic statistical tools for analyzing a model parameter, starting with the standard deviation (SD). Indeed SD, most commonly represented in Mathematics by the lower case Greek letter σ, provides an indication of how far the estimated parameters vary or "deviate" from the estimated mean values. In general, SD tells the researcher how spread out the model outputs are. Are they concentrated around the mean, or scattered far and wide? Basically, a small SD means that the values in a statistical data set are close to the mean of the data set, on average, and a large SD means that the values in the data set are farther away from the mean, on average. Taking this information into account, it is tempting to say that the estimated parameters with small SDs are the most sensitive, the most influential, the most important and deserve to be addressed with great care. This means, for example, just looking at the SDs of the estimated parameters shown in Figure 5.3D,E,F, that the parameter $R_{D_{22}}$ is important, sensitive. This is inconsistent with the starting hypothesis since the parameter $R_{D_{22}}$, considered important here by virtue of its small SD, has not fully converged and therefore, according to the hypothesis, should be considered as a parameter of little importance. However, there is no reason to be disappointed since in statistical studies where the parameters to be analyzed have very

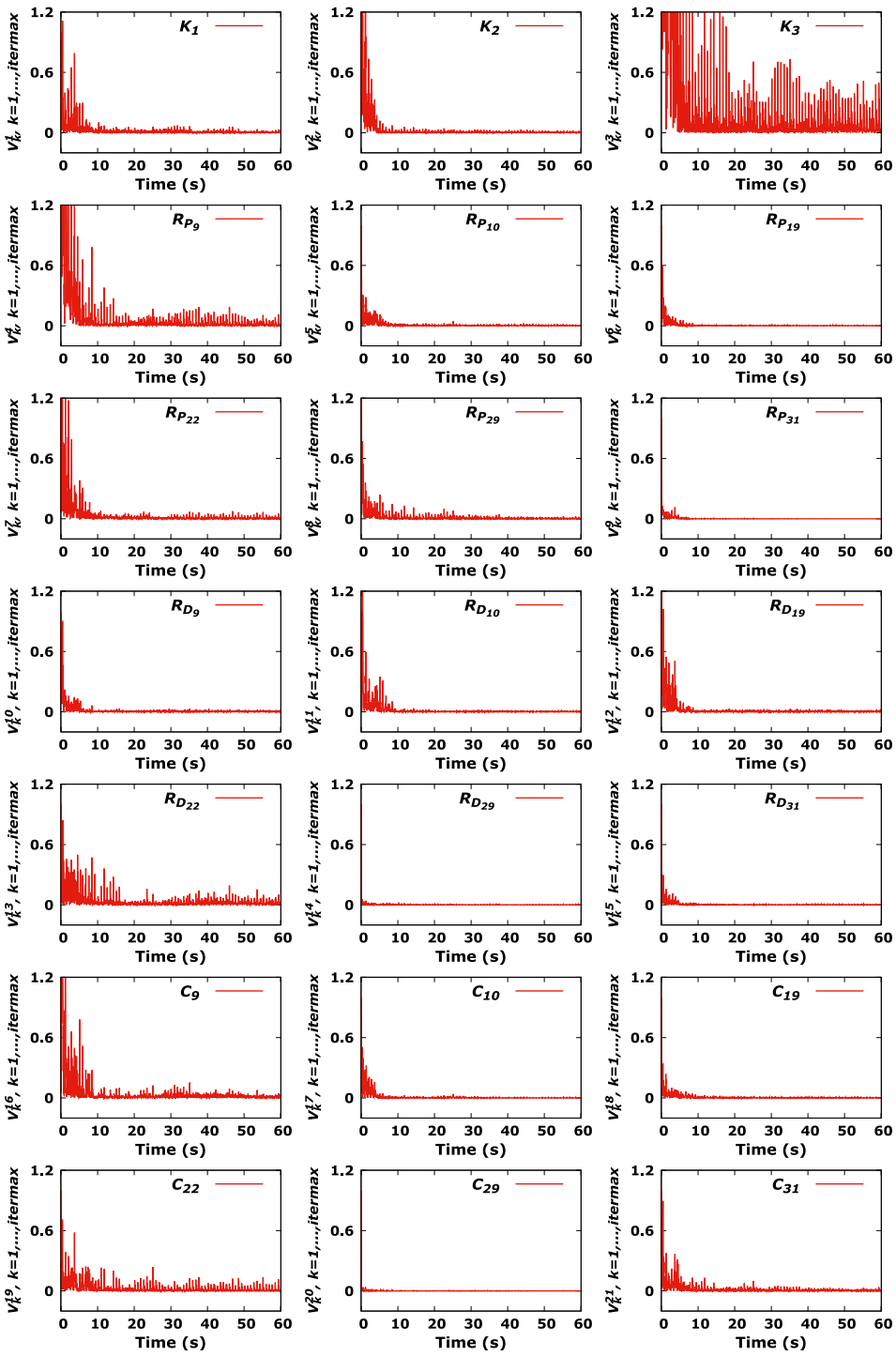

Figure 5.5: Variability of the 21 previously estimated blood flow model parameters [277] during the EnKF-based algorithm iterations. The maximum number of iterations is referred as *itermax*. Parameters with too much variability along all the EnKF-based algorithm iterations did not fully converge.

disparate, non-homogeneous scales of values as is the case here (the ratio between extreme values of estimated parameters is about 10^{20}), SD alone is most often not a relevant indicator of the dispersion of values around the mean. SD is therefore not discriminating enough and no clear conclusion based on it can be drawn about the base-case. In general, the coefficient of variation, which provides a general index, independent of the units of measurement employed, is the reference statistical tool to better deal with these cases.

5.4.3 Hypothesis testing using the coefficient of variation

As discussed above, a comparison of the estimated parameters importance on the basis of SDs is only relevant when comparing statistical series or observed distributions without extreme values and with means of the same order of magnitude. When the means have different orders of magnitude as is the case of estimated parameters here, it is better to consider the coefficient of variation (CV) rather than the simple SD. Indeed, unlike the value of SD, the actual value of the CV is independent of the unit in which the measurement was performed, so it is a dimensionless number. Let C_v be the mathematical expression of the CV. Then C_v, often expressed as a percentage, is defined as the ratio of the SD, σ, to the mean, μ. Its absolute value is referred to as relative standard deviation (RSD) [344]. This means that

$$C_v = \frac{\sigma}{\mu} \quad \text{and} \quad \text{RSD} = |C_v|.$$

As a standardized relative measure of the dispersion of statistical data around the mean, CV is useful since it shows the extent of variability with respect to the mean of a statistical population and is very easy to interpret. Indeed, higher the CV (respectively the RSD), further the values are from the mean and lower the CV (respectively the RSD), closer the values are from the mean. Again, the natural tendency would be to say that the lower the RSD of a given estimated parameter, the more influential that parameter is, important for estimating blood flow rate, and therefore, according to the initial hypothesis, has fully converged. And conversely, the higher the RSD of a given

estimated parameter is, the more non-influential that parameter is, unimportant for the estimation of blood flow, and could not accordingly fully converge.

Figures 5.6 and 5.7 show the time evolution of the RSD for the 21 parameters previously estimated in [277] and the mean values of the RSD for each of them, respectively. On closer examination, it can be observed that there is no obvious link between the RSDs and the non-full convergence of some of the estimated parameters. Indeed, the RSDs can explain some of the behavior (convergence) for estimated parameters but some variables still need more investigation. For instance, C_{22} and most of R_Ds (except $R_{D_{22}}$) meet the criterion intuited above for the interpretation of the RSDs for the estimated parameters, but in contrast C_{10}, C_{19}, C_{29}, and $R_{D_{22}}$ do not meet this criterion. The RSD for $R_{D_{22}}$ is low while $R_{D_{22}}$ has not fully reached a converged status as shown by its evolution and its high variability over time in Figures 5.3E and 5.5, respectively. In contrast, the RSD for C_{19} is high while C_{19} has fully converged as shown by its evolutions and its low variability over time over in Figures 5.3F and 5.5, respectively. This indicates that the RSDs are also not enough discriminating to properly confirm or disprove the initial hypothesis. However, one advantage in the use of RSDs is that they do not require any extra calculations. This is important since we would like to find a stopping criterion only using available information. In the next Section, a much more sophisticated statistical tool for sensitivity analysis (SA), namely the Sobol' indices, will be used to quantify the influence of the estimated parameters on the observed blood flow rate estimate. But before that, we would like to recall the state of the art in SA.

5.5 The state of the art in sensitivity analysis

The state of the art in SA has advanced significantly in the last decades. Indeed, with the development, construction or use of mathematical models in the scientific research, SA has become a valuable tool in the research world because it addresses several issues, including issues of fidelity and reliability of the model studied with respect

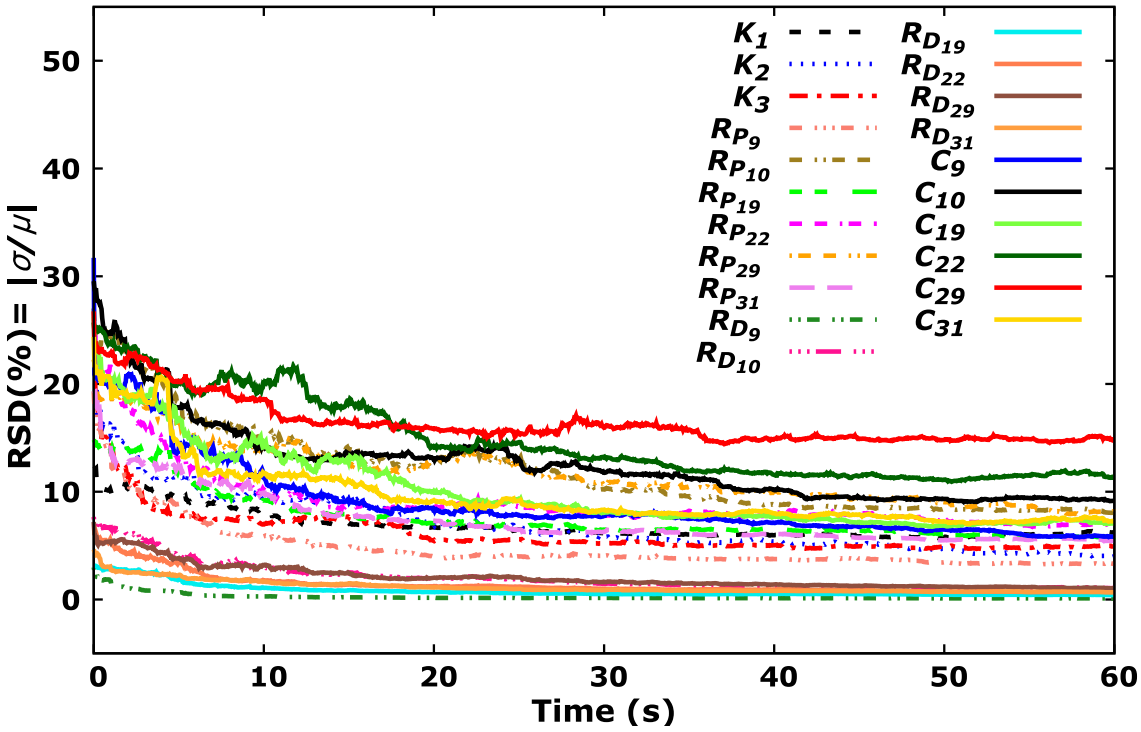

Figure 5.6: Time evolution of the relative standard deviations (RSDs) for the 21 parameters previously estimated in [277] using the EnKF-based parameter estimation algorithm.

to the modelled phenomenon, questions of influence or not of the input variables to the variability of the model output and also issues of interaction or not between input variables or groups of input variables. In fact, when we talk about SA, two important concepts emerge: sensitivity and contribution. Sensitivity, e.g. $\frac{\partial \mathbf{Y}}{\partial \mathbf{X}_i}$ (where $\mathbf{Y}$ represents a model output and $\mathbf{X}_i$, the i^{th} component of the model input variables, $\mathbf{X}$), gives an idea of how the output can respond to potential variations in input parameters. Contribution is the product of sensitivity and importance, e.g. $\frac{\partial \mathbf{Y}}{\partial \mathbf{X}_i} \times \sigma(\mathbf{X}_i)$. It is therefore used to determine the weight of an input variable (or group of input variables) on the output uncertainty.

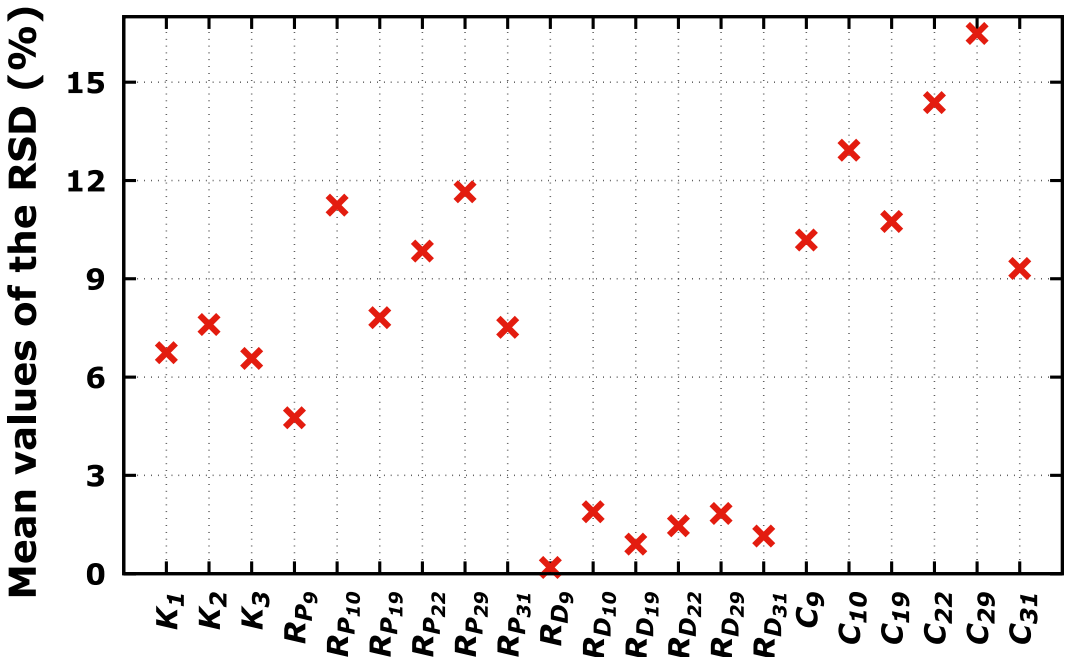

Figure 5.7: Average over all the EnKF-based algorithm iterations of the relative standard deviations (RSDs) for the 21 parameters previously estimated in [277] using the inversion procedure.

5.5.1 General settings

In general, the starting point of SA is a mathematical model that, to a set of random input variables $\mathbf{X}$, maps, via a function f, a random output/response variable $\mathbf{Y}$:

$$\begin{aligned} f: \quad \mathbb{R}^d &\longrightarrow \mathbb{R} \\ \mathbf{X} &\longmapsto \mathbf{Y} = f(\mathbf{X}), \end{aligned}$$

where f can be a very complex function (system of partial differential equations, system of ordinary differential equations, $\cdots$), and is in practice evaluated using a numerical solver, more or less expensive in computing time. The set of input variables $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_d)$ represents all the entities considered as random in the model.

SA studies how uncertainties in the input variables of a model (numerical or otherwise) cause uncertainties in the output variable [160, 166]. In other words, SA investigates how the variation in the output of a model can be attributed to variations of its input variables [336]. The reader is referred to Saltelli for more details [160, 175].

In general, SA is applied for multiple purposes, including method or code of calculation validation, research and development efforts orientation, justification in terms of dimensioning safety or system modification, and also as pointed out by Pianosi and many others, support model calibration, verification and simplification, support the prioritization of efforts for uncertainty reduction, or help with model-based decision-making [336, 345, 346, 347]. Such purposes are generally implemented as four main objectives of SA [160, 348]:

- the Factors Fixing (FF) setting or screening, which answers the question what input model variables have negligible effects on the model output so that the model should be rewritten without them [345]? This means that FF aims at identifying unessential model input variables, i.e., the ones, if any, which do not give any contribution to the model output uncertainty and therefore can be fixed to any value within their feasible range without a significant implication on the model output uncertainty [336, 349]. For this reason, FF is in practice used as a tool to reduce the number of uncertain input variables by fixing unessential factors. It is also used as a preliminary step before the calibration of model inputs using some available information (real output observations, constraints, etc.) [348];

- the Factors Prioritization (FP) setting or ranking aims at identifying the most important model input variables, (i.e., the ones that, if fixed, would lead to the greatest reduction in the uncertainty of the output) and at ranking them according to their relative contribution to the model output uncertainty [350, 351, 352]. In other words, FP answers the question in what order of importance do the model input variables influence the model outputs?

- The Variance Cutting (VC) setting is used to reduce the output variance from its initial value to a lower prescribed threshold value. VC aims at obtaining specific sensitivities for the different input factors and is, for example, applied in reliability and risk assessment studies [342, 348];

- the Factors Mapping (FM) aims at identifying the model input variables that produce critical model output values. In other words, FM aims at determining the important inputs in a specific domain of the output values, for instance which combination of factors produces output values above or below a pre-established threshold [348, 353, 354].

In general, these SA objectives and methods in turn can be gathered into two main categories [348]:

- The first one is Local Sensitivity Analysis (LSA) where the variability of the model output is explored around some specific input factor (the mean of a random variable, for instance) [355, 356, 357]. In other words, the LSA methods quantitatively evaluate the impact of a small variation around a given value of the inputs. This deterministic approach is therefore interested in calculating or estimating some kind of derivative or partial derivatives (gradient) of the model at a specific point of the input variable space [348, 355, 358, 359].

- The second one is Global Sensitivity Analysis (GSA), which unlike LSA rather attempts to explore the entire space of the input factors. This means that the GSA methods are interested in the variability of the model output in the entirety of domain of possible input parameter variations [341, 358, 360]. Therefore, GSA is useful for FP and FF. Indeed GSA provides indices that quantify the relative contribution to output uncertainty from each input variable. A sensitivity index of zero means that the associated input is non-influential (which is useful for FF) while the higher the index the more influential the input variable (FP). The Sobol' indices that will be used in this work belong to this category.

5.5.2 The Sobol' sensitivity indices

The literature on sensitivity measures is abundant [166, 175, 179, 358, 361, 362, 363, 364, 365]. According to Sudret, among this vast literature on sensitivity measures,

the Sobol' sensitivity indices have received much attention since they provide accurate information for most models [359]. Indeed, first introduced by Sobol in 1990 [178] and based on his earlier work on the Fourier Haar series [366], the indices were developed for the purpose of estimating not only the functional structure of the model output $\mathbf{Y}$ but also the sensitivity of $\mathbf{Y}$ with respect to any model input variable or any subgroup of model input variables. In this Section, we will closely follow the excellent papers on SA by Sobol [179] and by Archer, Saltelli and Sobol [361] in order to recall the essential features of the Sobol' sensitivity indices method.

5.5.2.1 Expansion into summands of increasing dimensions (ANOVA-representation)

Let $\mathbf{I}$ the unit interval $[0,1]$, $\mathbf{I}^d = \{\mathbf{X} : 0 \leq \mathbf{X}_i \leq 1,\, i = 1, \cdots, d\}$ the d-dimensional unit hypercube, and $i_1, \cdots, i_s$ where $1 \leq i_1 < \cdots < i_s \leq d$ and $s = 1, \cdots, d$, a group of indices. Given a mathematical model described by a square-integrable function f defined in $\mathbf{I}^d$ as follows:

$$\begin{array}{rccl} f: & \mathbf{I}^d & \longrightarrow & \mathbb{R}^p \\ & \mathbf{X} & \longmapsto & \mathbf{Y} = f(\mathbf{X}) = f(\mathbf{X}_1, \cdots, \mathbf{X}_d). \end{array}$$

According to Sobol, the function f can be uniquely expanded in a series of summands of increasing dimensions:

$$f(\mathbf{X}) = f_0 + \sum_{s=1}^{d} \sum_{i_1 < \cdots < i_s}^{d} f_{i_1 \cdots i_s}(\mathbf{X}_{i_1}, \cdots, \mathbf{X}_{i_s}), \tag{5.4}$$

provided that f_0 is constant and the integrals of the summands $f_{i_1 \cdots i_s}$ over any of their own variables are zero, i.e.,

$$\int_{\mathbf{I}} f_{i_1 \cdots i_s}(\mathbf{X}_{i_1}, \cdots, \mathbf{X}_{i_s})\, \mathrm{d}\mathbf{X}_{i_k} = 0, \quad 1 \leq k \leq s. \tag{5.5}$$

Formula (5.4) is also known as the ANOVA (term coming from ANalysis Of VAriances) decomposition of $f(\mathbf{X})$ and can also be written as

$$f(\mathbf{X}) = f_0 + \sum_{i=1}^{d} f_i(\mathbf{X}_i) + \sum_{1\leq i<j\leq d} f_{ij}(\mathbf{X}_i, \mathbf{X}_j) + \cdots + f_{12\cdots d}\,(\mathbf{X}_1, \cdots, \mathbf{X}_d). \tag{5.6}$$

It follows from the expressions (5.4) and (5.5) these important properties of the ANOVA decomposition:

- Formula (5.4) contains exactly $1 + \sum_{j=1}^{d} \binom{d}{j} = 2^d$ summands.
- $$f_0 = \int_{\mathbf{I}^d} f(\mathbf{X})\ \mathrm{d}\mathbf{X}. \tag{5.7}$$
- All the summands on the right-hand side of (5.4) are orthogonal meaning that for any two different summands $f_{i_1\cdots i_s}$ and $f_{j_1\cdots j_l}$ and for any group of indices $\{i_1 \cdots i_s\} \neq \{j_1 \cdots j_l\}$

  $$\int_{\mathbf{I}^d} f_{i_1\cdots i_s}(\mathbf{X}_{i_1}, \cdots, \mathbf{X}_{i_s}) f_{j_1\cdots j_s}(\mathbf{X}_{j_1}, \cdots, \mathbf{X}_{j_l})\ \mathrm{d}\mathbf{X} = 0. \tag{5.8}$$

- Any summand $f_{i_1\cdots i_s}(\mathbf{X}_{i_1}, \cdots, \mathbf{X}_{i_s})$ can be expressed as a multi-dimensional integral of $f(\mathbf{X})$:

  - $f_0 = \int_{\mathbf{I}^d} f(\mathbf{X})\ \mathrm{d}\mathbf{X} = \mathbb{E}(\mathbf{Y})$
  - $f_i(\mathbf{X}_i) = \int_{\mathbf{I}} \cdots \int_{\mathbf{I}} f(\mathbf{X}) \prod_{k\neq i} \mathrm{d}\mathbf{X}_k - f_0 = \mathbb{E}(\mathbf{Y}|\mathbf{X}_i) - f_0 = \mathbb{E}(\mathbf{Y}|\mathbf{X}_i) - \mathbb{E}(\mathbf{Y})$
  - $$\begin{aligned} f_{ij}(\mathbf{X}_i, \mathbf{X}_j) &= \int_{\mathbf{I}} \cdots \int_{\mathbf{I}} f(\mathbf{X}) \prod_{k\neq i,j} \mathrm{d}\mathbf{X}_k - f_0 - f_i(\mathbf{X}_i) - f_j(\mathbf{X}_j) \\ &= \mathbb{E}(\mathbf{Y}|\mathbf{X}_i, \mathbf{X}_j) - \mathbb{E}(\mathbf{Y}|\mathbf{X}_i) - \mathbb{E}(\mathbf{Y}|\mathbf{X}_j) + f_0 \\ &= \mathbb{E}(\mathbf{Y}|\mathbf{X}_i, \mathbf{X}_j) - \mathbb{E}(\mathbf{Y}|\mathbf{X}_i) - \mathbb{E}(\mathbf{Y}|\mathbf{X}_j) + \mathbb{E}(\mathbf{Y}) \end{aligned}$$
  - $\vdots$

  and so on for the higher dimension summands.

5.5.2.2 Functional decomposition of variance and Sobol' indices estimation

Squaring the expansion (5.4) and integrating over $\mathbf{I}^d$ and using the orthogonality property (Eq. (5.8)), the total variance D of $f(\mathbf{X})$ can be obtained. Indeed, we get

$$\int_{\mathbf{I}^d} f^2(\mathbf{X})\ \mathrm{d}\mathbf{X} = f_0^2 + \sum_{s=1}^{d} \sum_{i_1<\cdots<i_s}^{d} \int_{\mathbf{I}^d} f_{i_1\cdots i_s}^2(\mathbf{X}_{i_1},\cdots,\mathbf{X}_{i_s})\ \mathrm{d}\mathbf{X}_{i_1}\cdots\mathrm{d}\mathbf{X}_{i_s}. \tag{5.9}$$

This means that

$$\int_{\mathbf{I}^d} f^2(\mathbf{X})\,\mathrm{d}\mathbf{X} - f_0^2 = \sum_{s=1}^{d} \sum_{i_1<\cdots<i_s}^{d} \int_{\mathbf{I}^d} f_{i_1\cdots i_s}^2(\mathbf{X}_{i_1},\cdots,\mathbf{X}_{i_s})\ \mathrm{d}\mathbf{X}_{i_1}\cdots\mathrm{d}\mathbf{X}_{i_s}.$$

Using Eq (5.7), this implies that

$$\int_{\mathbf{I}^d} f^2(\mathbf{X})\ \mathrm{d}\mathbf{X} - \left(\int_{\mathbf{I}^d} f(\mathbf{X})\,\mathrm{d}\mathbf{X}\right)^2 = \sum_{s=1}^{d} \sum_{i_1<\cdots<i_s}^{d} \int_{\mathbf{I}^d} f_{i_1\cdots i_s}^2(\mathbf{X}_{i_1},\cdots,\mathbf{X}_{i_s})\ \mathrm{d}\mathbf{X}_{i_1}\cdots\mathrm{d}\mathbf{X}_{i_s}$$

i.e.,

$$\begin{aligned} D &= \sum_{s=1}^{d} \sum_{i_1<\cdots<i_s}^{d} \int_{\mathbf{I}^d} f_{i_1\cdots i_s}^2(\mathbf{X}_{i_1},\cdots,\mathbf{X}_{i_s})\,\mathrm{d}\mathbf{X}_{i_1}\cdots\mathrm{d}\mathbf{X}_{i_s} \\ &= \sum_{s=1}^{d} \sum_{i_1<\cdots<i_s}^{d} D_{i_1\cdots i_s}, \end{aligned}$$

where $D_{i_1\cdots i_s}$ are the variances of $f_{i_1\cdots i_s}$. The fraction of the total variance of $f(\mathbf{X})$ which is due to any individual summand in Eq (5.4), referred as $S_{i_1\cdots i_s}$ is called global sensitivity indices and is given by the following relation:

$$S_{i_1\cdots i_s} = \frac{D_{i_1\cdots i_s}}{D}.$$

It follows that all the global sensitivity indices are nonnegative and their sum over all possible combinations of indices is

$$\sum_{s=1}^{d} \sum_{i_1<\cdots<i_s}^{d} S_{i_1\cdots i_s} = \frac{\displaystyle\sum_{s=1}^{d} \sum_{i_1<\cdots<i_s}^{d} D_{i_1\cdots i_s}}{D} = 1.$$

The Sobol' indices $S_{i_1\cdots i_s}$ thus calculated, are easy to interpret. As they are all nonnegative and their sum over all possible combinations of indices is 1, the larger the

index (close to 1), the more important the corresponding variable will be. Moreover, as the Sobol' indices measure the amount of the total variance D coming from the uncertainties in a set of input variables, the first order indices, S_i, quantify the influence of each variable taken separately and higher order indices account for the interactive contributions to the total variance D. For example, S_1 is the main effect of the variable $\mathbf{X}_1$. It expresses the sensitivity of the output $\mathbf{Y}$ with respect to the input variable $\mathbf{X}_1$, i.e., the part of the variance of $\mathbf{Y}$ due to $\mathbf{X}_1$. S_{13} is the interaction effect of variables $\mathbf{X}_1$ and $\mathbf{X}_3$. It denotes the sensitivity of the variance of $\mathbf{Y}$ with respect to the interaction of the variables $\mathbf{X}_1$ and $\mathbf{X}_3$, i.e., that part of the output variation due to variables $\mathbf{X}_1$, $\mathbf{X}_3$ which cannot be explained by the sum of the effect of the two variables alone. S_{123} expresses the sensitivity of the variance of $\mathbf{Y}$ with respect to the variables $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$ which is not taken into account in the effect of the variables alone and the interactions two by two. And so on until the last term $S_{12\cdots d}$, the index of order d, which is that fraction of the output variance which cannot be explained by summing terms of lower order. The number of sensitivity indices thus constructed, from first order to order d, is $2^d - 1$. Therefore, when the number of input variables d becomes too large, the number of sensitivity indices becomes exceedingly large so that estimating and interpreting all these indices quickly becomes impossible. It is for these reasons that the practitioner does not usually estimate indices of order higher than two. However, a possible alternative for situations where the input parameters are too large was introduced by Homma and Saltelli [358] with the total sensitivity indices.

5.5.2.3 The total sensitivity indices

Also called "total indices" or "total effects", these indices express the total sensitivity of the variance of $\mathbf{Y}$ with respect to any given input variable, i.e. the sensitivity with respect to this variable in all its forms (sensitivity with respect to the variable alone and sensitivity with respect to the interactions of this variable with other variables).

The total sensitivity index, S_{T_i}, with respect to the variable $\mathbf{X}_i$ is therefore defined as the sum of all sensitivity indices in regard to the variable $\mathbf{X}_i$ and can be written as follows [367]:

$$S_{T_i} = S_i + \sum_{i<j} S_{ij} + \sum_{j\neq i, k\neq i, j<k} S_{ijk} + \cdots = \sum_{l \in @i} S_l$$

where $@i$ represents all the subsets of $\{1, \cdots, d\}$ including i. For instance, for a model with three input variables, the total sensitivity index with respect to the variable $\mathbf{X}_1$ is $S_{T_1} = S_1 + S_{12} + S_{13} + S_{123}$.

However, it is worth noting that when the assumption of independence of model input factors is not satisfied (e.g., in input-dependent models), first-order indices no longer necessarily express sensitivity with respect to single input variables if the latter are correlated with other input variables. In this case, if these input variables can be partitioned into clusters of dependent variables (variables in a same cluster are dependent but variables in two different clusters are independent), an approach similar to that used to compute the classical Sobol' sensitivity indices can be adopted by considering multidimensional sensitivity indices. These express the sensitivity of the output variance with respect to one cluster of model input factors. For more details, the reader is referred to Jacques et al. [368].

5.5.2.4 Numerical estimation of the Sobol' sensitivity indices

For numerical computation of Sobol' sensitivity indices, several techniques and algorithms have been developed and can be found in the literature. For instance, the reader is referred to Sobol [178, 179] and Saltelli [369] where Monte Carlo sampling based methods are proposed either for first order and interaction indices computation [179] or for first order and total indices computation [369]. Given the cost of such methods in terms of number of model calls (the complexity of computing is $\mathcal{O}(N^{-1/2})$ where N represents the sample size) [367], many extensions of these methods and other alternative methods for reducing the complexity of indices

computation have been proposed by several researchers [176, 359, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381]. These proposed extensions include pseudo-probabilistic simulation methods such as Latin Hypercube Sampling (LHS), Fourier Amplitude Sensitivity Test (FAST) based methods, Local polynomial and polynomial chaos expansions based methods, Quasi-Monte Carlo (QMC) based methods, etc., and their variants. The details of these different methods are beyond the scope of this paper. Indeed, in the present work, the package sensitivity of `R` environment available on the CRAN website[1] [382] will be used to compute the Sobol' sensitivity indices. The objective will be to estimate both first order and total indices in order to be able to take into account the interactions of each parameter with all the others. Thus, the function `sobolmartinez` of the package sensitivity, which implements the Monte Carlo estimation of the Sobol' indices for both first-order and total indices, will be used.

5.6 Application to an analytical model and a patient-specific blood flow model

The effectiveness of the Sobol' indices is tested using two distinct models. The objective is to put into practice the Sobol' indices theory described above by identifying on two concrete examples, the most influential model input parameters. Before the performance of the Sobol' indices is tested on a patient-specific clinical data, a test case using an analytical model is carried out. This case is presented in order to show the capability of the Sobol' indices on simple cases.

5.6.1 Illustration on an analytical function

The efficiency of the Sobol' indices is investigated using an analytical function. Let us consider a mathematical model described by a square-integrable function f defined as follows:

[1] `https://cran.r-project.org/web/packages/sensitivity/sensitivity.pdf`

$$\begin{aligned} f: \ [-\tfrac{1}{2},\tfrac{1}{2}]^3 &\longrightarrow \mathbb{R} \\ \mathbf{X} &\longmapsto \mathbf{Y} = f(\mathbf{X}) = f(\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3) = 5\mathbf{X}_1^2 + 2\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_3, \end{aligned}$$

where $\mathbf{X}_i$, $i = 1,2,3$ are independent variables and $\mathbf{X}_i \sim U[-\frac{1}{2},\frac{1}{2}]$ for $i = 1,2,3$, meaning that the probability density functions (PDFs) of its input parameters are assumed as follows:

$$p_i(\mathbf{X}_i) = \begin{cases} 1 & \text{if } -\frac{1}{2} \le \mathbf{X}_i \le \frac{1}{2} \\ & \qquad\qquad \text{for } i = 1,2,3 \\ 0 & \text{if } \mathbf{X}_i < -\frac{1}{2} \text{ or if } \mathbf{X}_i > \frac{1}{2} \end{cases}$$

Following the Sobol decomposition theory presented in Section 5.5.2.1 (the reader can easily verify that this function f meets the assumptions set out by Sobol), f can be expanded in a series of 8 summands as below:

$$\begin{aligned} f(\mathbf{X}) &= f_0 + f_1(\mathbf{X}_1) + f_2(\mathbf{X}_2) + f_3(\mathbf{X}_3) + f_{12}\left(\mathbf{X}_1,\mathbf{X}_2\right) + f_{13}\left(\mathbf{X}_1,\mathbf{X}_3\right) + f_{23}\left(\mathbf{X}_2,\mathbf{X}_3\right) \\ &+ f_{123}\left(\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3\right), \end{aligned}$$

where any summand is computed as follows:

- $$\begin{aligned} f_0 = \mathbb{E}(\mathbf{Y}) &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} f(\mathbf{X})\,\mathrm{d}\mathbf{X} \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(5\mathbf{X}_1^2 + 2\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_3\right)\mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2\,\mathrm{d}\mathbf{X}_3 = \frac{5}{12}, \end{aligned}$$
- $$\begin{aligned} f_1(\mathbf{X}_1) = \mathbb{E}(\mathbf{Y}|\mathbf{X}_1) - f_0 &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} f(\mathbf{X})\,\mathrm{d}\mathbf{X}_2\,\mathrm{d}\mathbf{X}_3 - f_0 \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(5\mathbf{X}_1^2 + 2\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_3\right)\mathrm{d}\mathbf{X}_2\,\mathrm{d}\mathbf{X}_3 - f_0 = 5\mathbf{X}_1^2 - \frac{5}{12}, \end{aligned}$$
- $$\begin{aligned} f_2(\mathbf{X}_2) = \mathbb{E}(\mathbf{Y}|\mathbf{X}_2) - f_0 &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} f(\mathbf{X})\ \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_3 - f_0 \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(5\mathbf{X}_1^2 + 2\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_3\right)\ \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_3 - f_0 = 2\mathbf{X}_2, \end{aligned}$$
- $$\begin{aligned} f_3(\mathbf{X}_3) = \mathbb{E}(\mathbf{Y}|\mathbf{X}_3) - f_0 &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} f(\mathbf{X})\ \mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2 - f_0 \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}} \left(5\mathbf{X}_1^2 + 2\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_3\right)\mathrm{d}\mathbf{X}_1\,\mathrm{d}\mathbf{X}_2 - f_0 = 0, \end{aligned}$$

- $f_{12}(\mathbf{X}_1, \mathbf{X}_2) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\mathbf{X})\, \mathrm{d}\mathbf{X}_3 - f_0 - f_1(\mathbf{X}_1) - f_2(\mathbf{X}_2)$
$= \mathbb{E}(\mathbf{Y}|\mathbf{X}_1, \mathbf{X}_2) - f_0 - f_1(\mathbf{X}_1) - f_2(\mathbf{X}_2)$
$= 5\mathbf{X}_1^2 + 2\mathbf{X}_2 - \frac{5}{12} - (5\mathbf{X}_1^2 - \frac{5}{12}) - 2\mathbf{X}_2 = 0,$

- $f_{13}(\mathbf{X}_1, \mathbf{X}_3) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\mathbf{X})\, \mathrm{d}\mathbf{X}_2 - f_0 - f_1(\mathbf{X}_1) - f_3(\mathbf{X}_3)$
$= \mathbb{E}(\mathbf{Y}|\mathbf{X}_1, \mathbf{X}_3) - f_0 - f_1(\mathbf{X}_1) - f_3(\mathbf{X}_3)$
$= 5\mathbf{X}_1^2 + \mathbf{X}_1\mathbf{X}_3 - \frac{5}{12} - (5\mathbf{X}_1^2 - \frac{5}{12}) - 0 = \mathbf{X}_1\mathbf{X}_3,$

- $f_{23}(\mathbf{X}_2, \mathbf{X}_3) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(\mathbf{X})\ \mathrm{d}\mathbf{X}_1 - f_0 - f_2(\mathbf{X}_2) - f_3(\mathbf{X}_3)$
$= \mathbb{E}(\mathbf{Y}|\mathbf{X}_2, \mathbf{X}_3) - f_0 - f_2(\mathbf{X}_2) - f_3(\mathbf{X}_3)$
$= 2\mathbf{X}_2 + \frac{5}{12} - \frac{5}{12} - 2\mathbf{X}_2 - 0 = 0,$

- $f_{123}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) = f(\mathbf{X}) - f_0 - f_1(\mathbf{X}_1) - f_2(\mathbf{X}_2) - f_3(\mathbf{X}_3) - f_{12}(\mathbf{X}_1, \mathbf{X}_2)$
$-f_{13}(\mathbf{X}_1, \mathbf{X}_3) - f_{23}(\mathbf{X}_2, \mathbf{X}_3)$
$= 5\mathbf{X}_1^2 + 2\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_3 - \frac{5}{12} - (5\mathbf{X}_1^2 - \frac{5}{12}) - 2\mathbf{X}_2 - \mathbf{X}_1\mathbf{X}_3 = 0.$

Consequently, the total variance D of $f(\mathbf{X})$ and the partial variances D_1, D_2, D_3, D_{12}, D_{13}, D_{23}, and D_{123} of $f_1(\mathbf{X}_1)$, $f_2(\mathbf{X}_2)$, $f_3(\mathbf{X}_3)$, $f_{12}(\mathbf{X}_1, \mathbf{X}_2)$, $f_{13}(\mathbf{X}_1, \mathbf{X}_3)$, $f_{23}(\mathbf{X}_2, \mathbf{X}_3)$, and $f_{123}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$, respectively, are analytically computed as:

- $D = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} f^2(\mathbf{X})\ \mathrm{d}\mathbf{X} - f_0^2$
$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(5\mathbf{X}_1^2 + 2\mathbf{X}_2 + \mathbf{X}_1\mathbf{X}_3\right)^2 \mathrm{d}\mathbf{X}_1\, \mathrm{d}\mathbf{X}_2\, \mathrm{d}\mathbf{X}_3 - f_0^2$
$= \frac{47}{72} - \left(\frac{5}{12}\right)^2 = \frac{23}{48},$

- $D_1 = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_1^2(\mathbf{X}_1)\ \mathrm{d}\mathbf{X}_1$
$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(5\mathbf{X}_1^2 - \frac{5}{12}\right)^2 \mathrm{d}\mathbf{X}_1 = \frac{5}{36},$

- $D_2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_2^2(\mathbf{X}_2)\ \mathrm{d}\mathbf{X}_2$
$= \int_{-\frac{1}{2}}^{\frac{1}{2}} (2\mathbf{X}_2)^2\ \mathrm{d}\mathbf{X}_2 = \frac{1}{3},$

- $D_3 = D_{12} = D_{23} = D_{123} = 0,$

- $D_{13} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} f_{13}^2(\mathbf{X}_1, \mathbf{X}_3)\ \mathrm{d}\mathbf{X}_1\ \mathrm{d}\mathbf{X}_3$
  $= \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} (\mathbf{X}_1\mathbf{X}_3)^2\ \mathrm{d}\mathbf{X}_1\ \mathrm{d}\mathbf{X}_3 = \frac{1}{144}.$

This implies the analytically calculated Sobol' sensitivity indices which are listed below in Table 5.1. These indices are compared with the ones computed using the function `sobolmartinez` from the package sensitivity of the `R` environment using the following arguments: the analytical function, f, and two 1,000-random samples generated following a uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$. The results, illustrated in Figure 5.8, show a very good agreement between the numerical method and the analytical one. This is both interesting and encouraging, especially when SA is performed on a model with large input parameters where the analytical determination of the Sobol' indices becomes almost impossible.

Table 5.1: The Sobol' indices analytically computed for the analytical example.

Variables	**Main effects**	**Total effects**	**Interaction effects**
$\mathbf{X}_1$	0.2898	0.3043	–
$\mathbf{X}_2$	0.6956	0.6956	–
$\mathbf{X}_3$	0	0.0145	–
$\mathbf{X}_1, \mathbf{X}_2$	–	0	0
$\mathbf{X}_1, \mathbf{X}_3$	–	–	0.0145
$\mathbf{X}_2, \mathbf{X}_3$	–	–	0
$\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$	–	–	0

Furthermore, as it is important to be able to give meaning to the Sobol' indices calculated, the following interpretation of the indices for the 3 parameters listed in Table 5.1 is suggested:

- $\mathbf{X}_2$, with a main effect value equivalent to the total effect one (about 0.6956), appears to be the most influential variable on the output variance, whether within the main or total indices framework. This means, on the one hand, that nearly 69.56% of the variance of $\mathbf{Y}$ is explained by $\mathbf{X}_2$ and, on the other hand, that $\mathbf{X}_2$ operates alone on the variance of $\mathbf{Y}$.

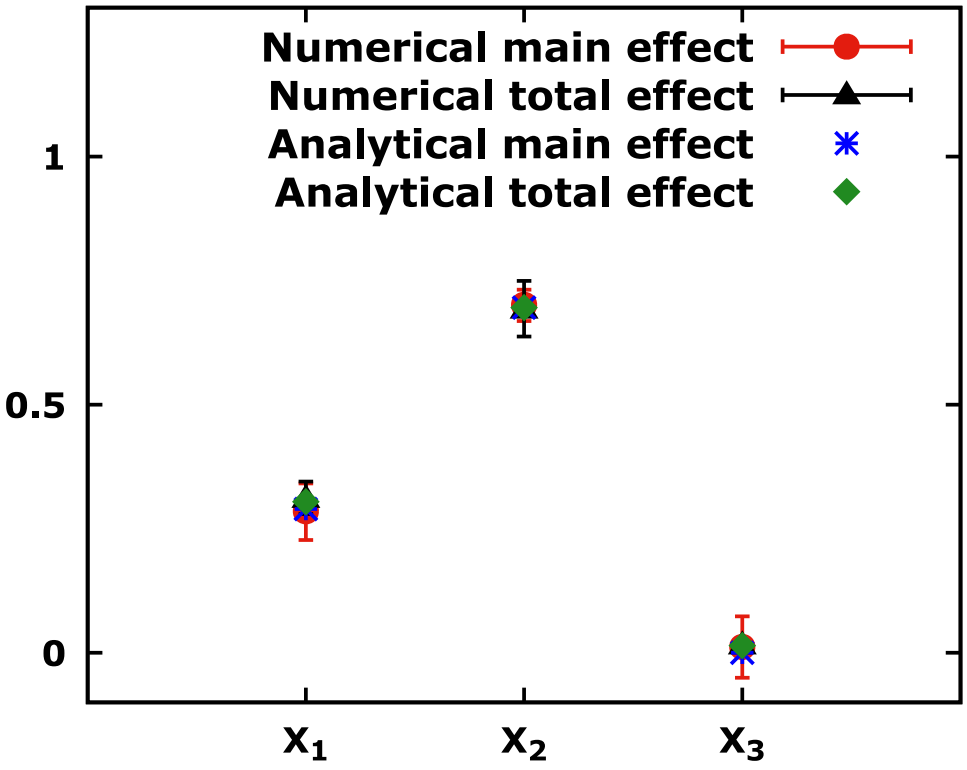

Figure 5.8: Comparative overview of the main and total indices calculated in two different ways for each of the three parameters of the analytical model. The numerically computed indices are in good agreement with the analytically computed ones.

- The variable $\mathbf{X}_1$, with a main effect of about 0.2898 has a relatively important influence alone, explaining nearly 28.98% of the variance of $\mathbf{Y}$ but remains much less influential than $\mathbf{X}_2$.

- The variable $\mathbf{X}_3$ has no influence on the output variance alone, but has a relatively small influence in interaction (with $\mathbf{X}_1$), with a total effect of about 0.0145.

- There is no part of output variation due to the interaction effect either between $\mathbf{X}_1$ and $\mathbf{X}_2$ ($S_{12} = 0$), or between $\mathbf{X}_2$ and $\mathbf{X}_3$ ($S_{23} = 0$) but also between $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$ ($S_{123} = 0$).

- To conclude, this implies that the variance of $\mathbf{Y}$ is due for 69.56% to $\mathbf{X}_2$, 28.98% to $\mathbf{X}_1$ and 1.45% to the interaction between $\mathbf{X}_1$ and $\mathbf{X}_3$. Note also that in this

example, the interaction between $\mathbf{X}_1$ and $\mathbf{X}_3$ is due to a non-additive relationship between these two variables in the model expression.

This analytical example shows how the Sobol' indices based approach can be used to perform SA. In the next Section, this technique is applied to the inversion problem associated with the blood flow model presented above in Section 5.2.

5.6.2 Application to a 0D blood flow model using patient-specific data

The performance of the Sobol' indices is now investigated on the blood flow model described in Section 5.2 in order to give a satisfactory answer to the initial hypothesis formulated in Section 5.3 about the non-full convergence of some estimated parameters. But before that, let us recall the procedure for generating the patient-specific data.

5.6.2.1 Data generation

The patient-specific data used in the current study comes mainly from medical imaging. These are essentially magnetic resonance angiography and magnetic resonance imaging (MRA & MRI) acquisitions, and a 3D time of flight magnetic resonance angiography (3D-TOF-MRA) images provided by the Department of Neuroradiology at the *Centre Hospitalier Régional Universitaire de Montpellier (CHRU), Montpellier, France.* Exploring these images with dedicated softwares, hemodynamic data, namely blood flow rates in ascending aorta (AA) and in the right and left internal carotid arteries (R-ICA and L-ICA), and morphological data including geometric measurements of some arteries have been extracted. This also made it possible to build a patient-specific arterial network consisting of 33 arteries including the aorta, vertebral, carotid and brachial arteries together with the complete circle of Willis as shown in Figure 5.1. The missing geometric measurements of other

arteries were obtained from average data reported in the literature [284, 286]. For more details, the reader is referred to previous works [12, 277].

The goal of this operation was to solve the minimization problem (the inverse hemodynamic problem) described in Section 5.2.2. More precisely, this consisted in finding the right set of model parameters that would allow to reproduce the flow rates observed in the patient-specific's right and left internal carotid arteries (ICAs). These target parameters, 21 in number, consist of 18 parameters (R_{P_i}, R_{D_i} and C_i where $i = 9, 10, 19, 22, 29, 31$ denotes the compartment numbers in Figure 5.1A) used to define the eleven Windkessel WK3 model boundary conditions in Figure 5.1, and of 3 constants (k_1, k_2, k_3) defining the product of Young modulus and thickness of arteries by means of this empirical formula: $Eh = r(k_1\, e^{k_2 r} + k_3)$. For this purpose, an EnKF-based parameter estimation algorithm was developed and used. The latter was executed for 60 s (approximately 72 cardiac cycles) under assumptions. For more details either about these assumptions or about the parameter estimation algorithm and its execution, the reader is again referred to previous works [12, 277].

During the parameter estimation algorithm execution, as early presented in Section 5.3, one noticed that from 8.35 s, there was a very good agreement between the observed and predicted blood flow waveforms and also that iterating more with the inversion algorithm did not modify the recovery of the observed blood flow rates (see Figure 5.4). This means that after 8.35 s (about 10 cardiac cycles, or about 200 EnKF-based algorithm iterations), the right set of model parameters allowing to reproduce the blood flow rates observed in the ICAs was found and therefore the execution of the algorithm could be stopped. But, to ensure not to stop the algorithm too early, we let it run until the end (i.e. until *itermax*, the maximum number of algorithm iterations, arbitrary fixed as 1,440) and only saved the results for the last 20 iterations, namely the 21 estimated parameters with their associated uncertainties, the errors between the simulated and observed blood flow rates in the ICAs (the cost function, J, in Section 5.2.2), and also the variability of the simulated blood flow rates

with respect to each estimated parameter at every iteration. These are used to make up the patient-specific database and to build a simplified model of the patient-specific blood flow model in order to be able to properly perform SA.

5.6.2.2 Sensitivity analysis and hypothesis testing using the Sobol' indices

To identify both the most influential, the weakly influential, and the non-influential input parameters in the patient-specific blood flow model in order to clearly and definitively address the question of the reliability of the original assumption stated above in Section 5.3, SA using the Sobol' indices is performed on a simplified model of the patient-specific blood flow model. For this purpose, the 21 estimated input parameters are assumed to be independent variables and their relationship with the cost function, J, is assumed to be governed by a simple linear model. This simplified model is based on the saved data in the last 20 EnKF-based algorithm iterations and is defined for each of these iterations as follows:

$$\begin{aligned} \mathcal{L}: \quad \mathbb{R}^{21} &\longrightarrow \mathbb{R}^{2} \\ \mathbf{X} &\longmapsto \mathcal{L}(\mathbf{X}) = \mathbf{B}\mathbf{X}^{\top} = J = (J_{RICA}, J_{LICA})^{\top}, \end{aligned}$$

where $\mathbf{X}$, $\mathcal{L}$, and J stand for the 21 estimated parameters (also referred as $\mathbf{x}$ in Section 5.2.2), the linear mapping associated to the matrix $\mathbf{B}$, and the errors between the simulated and observed blow flow rates in R-ICA and L-ICA, respectively, in each of the last 20 EnKF-based algorithm iterations. The components of J, J_{RICA} and J_{LICA}, represent the cost function obtained when using only the blood flow rate in R-ICA as observed data and the cost function obtained when using only the blood flow rate in L-ICA as observed data, respectively, in each of the last 20 iterations. For each of these iterations, $\mathbf{B}$ is a 2-by-21 matrix obtained by assembling the local variability values of the components of the cost function, J, with respect to each estimated parameter. This means that for each iteration considered,

$$\mathbf{B} = \left(\frac{\partial J_{RICA}}{\partial \mathbf{X}_i}, \frac{\partial J_{LICA}}{\partial \mathbf{X}_i} \right)^{\top}, i = 1, \cdots, 21,$$

where $\frac{\partial J_{RICA}}{\partial \mathbf{X}_i}$ and $\frac{\partial J_{LICA}}{\partial \mathbf{X}_i}$ refer to the local variability values of the cost function, J_{RICA}, and to the local variability values of the cost function, J_{LICA}, respectively, with respect to each estimated parameter.

The Sobol' first order indices and total indices for the 21 estimated parameters in each of the last 20 iterations are estimated using the same function from the `R` package sensitivity as in the analytical example with the following arguments: the simplified model defined above and two 10,000-random samples generated following a multivariate normal distribution with means equal to mean values of the 21 estimated parameters at each iteration considered and covariance matrix whose the elements consist of the standard deviations associated to the 21 estimated parameters in each iteration.

Typical results of the first order and total indices, computed for some of the last 20 EnKF-based algorithm iterations, are shown in Figure 5.9 when using only either J_{RICA} or J_{LICA} as the simplified model output. The interpretation of these results is not entirely straightforward, obvious. It is even very delicate and should therefore be done with care. Indeed, from one Figure to another, diverse and various patterns are visible. However, looking very carefully and closely at these Figures, a general trend emerges. Indeed, the blood flow model behaves as if it depends upon only a few parameters, with the rest having comparatively little influence. More precisely, only five of the 21 estimated parameters appear to be important. These are all proximal resistance parameters ($R_{P_{10}}$, $R_{P_{19}}$, $R_{P_{22}}$, $R_{P_{29}}$, $R_{P_{31}}$), with $R_{P_{31}}$ in the lead. This means that $R_{P_{31}}$ seems to be the most influential parameter, $R_{P_{10}}$, $R_{P_{19}}$, $R_{P_{22}}$, and $R_{P_{29}}$, the slightly influential ones and all the others seem to be non-influential. This is very surprising since it is rather the distal resistances that usually appear to be the most important parameters in most cases. In order to review this general trend and to provide for each estimated parameter single representative main and total indices, which are much more easy to interpret, the average of the first order and total indices over all the last 20 EnKF-based algorithm iterations for each of the 21 estimated parameter has been calculated. The results are shown in Figure 5.10 when using either only

J_{LICA} (Figure 5.10A) or only J_{RICA} (Figure 5.10B) as the simplified model output. The results from these Figures are perfectly consistent with the general trend identified above and therefore confirm it. Indeed, from the main and total mean indices reported in Figure 5.10, five indices measuring the importance of the five proximal resistances cited above as influential parameters stand out from all the others with the same order of importance. Comparing these results with those illustrated in Figures 5.3 and 5.5, it can be seen that the parameters found not to be fully converged in these Figures have been identified here as non-influential parameters. Furthermore, it is also noted that all the parameters identified here as influential had fully converged, as shown in Figures 5.3 and 5.5. This is consistent with the hypothesis formulated from the outset and therefore, confirms and validates it at the same time.

However, it is still legitimate to ask the question of the reliability, the relevance or the credibility of these results. To ensure that these results about the importance of estimated parameters make sense, the factors fixing setting and the factors prioritization setting are performed. In concrete terms, the EnKF-based parameter estimation algorithm is executed using the following two scenarii with regard to the set of model input parameters:

- **Scenario A**: The algorithm is executed with a set of input parameters created by keeping unchanged the default initialization values of the five important parameters detected above, and fixing during all the EnKF-based algorithm iterations all other parameters to their nominal values taken in previous works [277].

- **Scenario B**: The algorithm is now executed using a set of input parameters created by fixing during all the EnKF-based algorithm iterations the five important factors detected above to their nominal values (estimated values) taken in previous works [277] and keeping unchanged the default initialization values of all other parameters.

At the end, for each of both scenarii considered, the normalized cost function over time, J_{norm}, which is obtained by dividing the cost function at each iteration, referred to as J by iter., by the cost function at the first iteration, referred to as J of 1^{st} iter., is compared to the reference normalized cost function i.e., that previously obtained using the set of the default input parameters [277]. The results of this comparison are shown in Figure 5.11. One can notice that there is virtually no significant change on the reference normalized cost function J_{norm} in scenario A. However, a reduction of the model output uncertainty can be observed in scenario B. Indeed, the values of the normalized cost function over time in scenario A and the reference ones are almost identical while those obtained in scenario B are relatively lower than the reference ones. This corroborates, on the one hand, the fact that the input parameters fixed in scenario A are really unessential in the blood flow rates estimation in the patient-specific ICAs, and at the same time, on the other hand, the fact that only the input parameters fixed in scenario B are important for the blood flow rates estimation in the patient-specific ICAs. This means that for the purpose of the blood flow rate estimation in this patient-specific ICAs, the blood flow model can be reduced to the five input parameters detected as influential instead of the twenty-one currently used while still performing well.

5.6.2.3 Choice of a stopping criterion in the inversion procedure

With the dust settled over the parameter importance issue, the essential question remains as to how a stopping criterion in the EnKF-based parameter estimation algorithm can be fixed in order to prevent from unnecessary over-solving. Many possibilities can be considered. However, in view of the whole approach followed so far, it seems judicious to propose for the blood flow model examined here, a stopping criterion based only on the variability of the most important parameters. Thus, the following stopping criterion is proposed: the EnKF-based algorithm execution is automatically interrupted as soon as the variabilities of the 5 important parameters at a given iteration k, V_k^i, $\quad i = 5, \cdots, 9$ in Figure 5.5, reach the threshold value of 10^{-4}.

The choice of this threshold value is based on the currently available values of the variabilities of the estimated parameters after 8.35 s (about 200 EnKF-based algorithm iterations). This typically leads to a stopping time $t = 10\,\mathrm{s}$, reducing by a factor of six the time to solution. The main advantage of establishing this stopping criterion is that it can be used to achieve faster diagnosis in clinical situations.

5.7 Concluding remarks

The usefulness of global sensitivity analysis in the identification and prioritization of the parameters importance in a patient-specific cardiovascular model has been demonstrated. The approach, based on a 0D blood flow model of the upper body arteries and of the circle of Willis, complements our previous works on uncertainty quantification and propagation and on backward sensitivity analysis. It has been shown how statistical indicators such as the variability of the model parameters over time, their standard deviations, their coefficients of variation and the Sobol 'indices can be used to both identify the model input parameters, responsible or not for the variability on the model output variance, and to establish a stopping criterion in the parameter estimation algorithm. The input parameters importance ranking based on the sensitivity measures demonstrated that the most influential factors involve most of the proximal resistances, largely responsible for the uncertainty in predicting the patient-specific blood flow rates in the left and right internal carotid arteries.

Overall, the global sensitivity analysis proved to be a powerful tool for explaining and quantifying uncertainties as well as providing insight into devising useful ways for reducing uncertainties in the blood flow rate estimation performance. Thus this study made it possible, in addition to the inversion results, to give an indication of the importance of each optimization parameter and also find when it is safe to stop the inversion procedure giving increasing confidence level on the outcome of the inversion, and this even in situations where some variables might have not fully converged to steady values during the inversion.

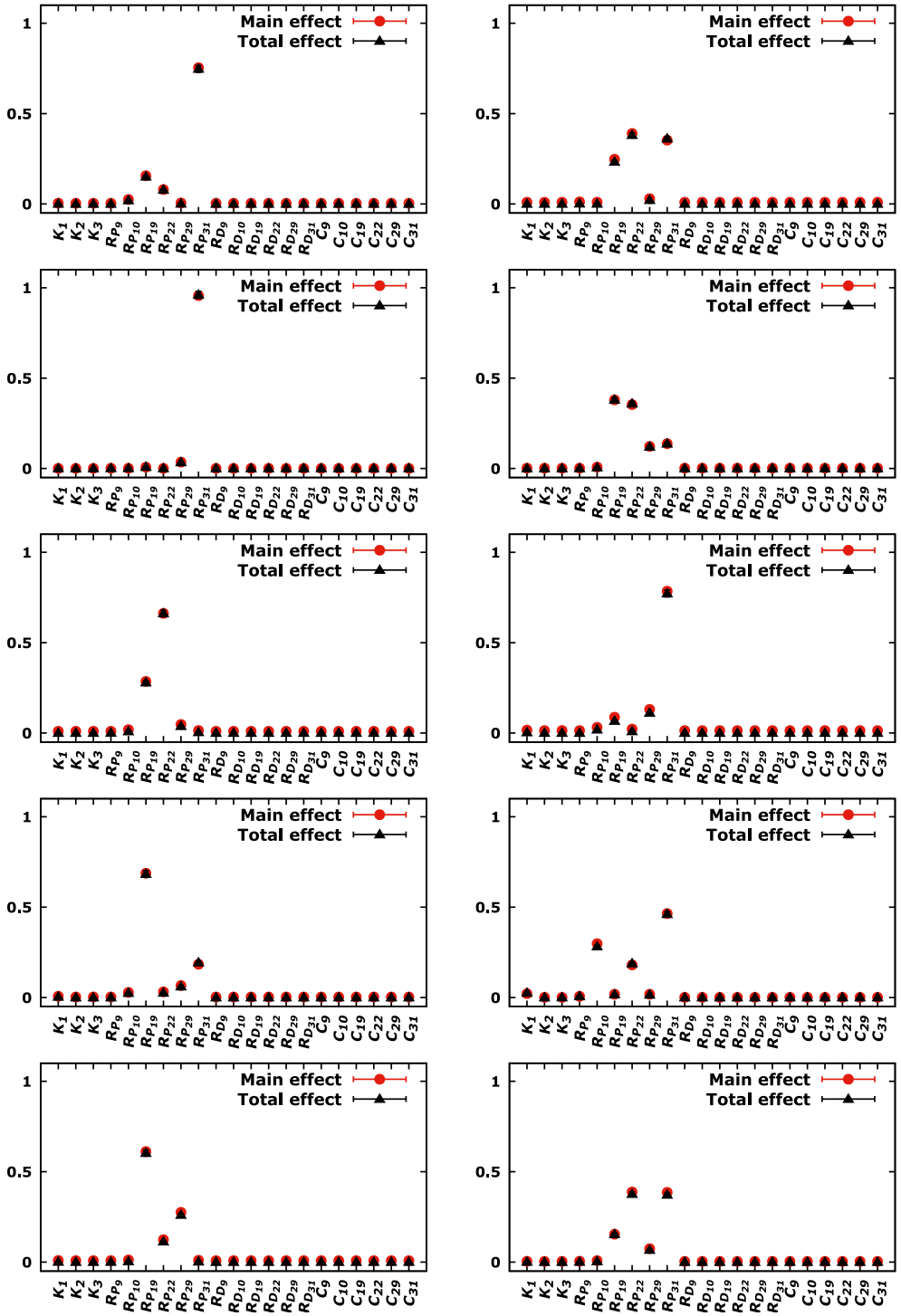

Figure 5.9: First order and total indices for the 21 estimated parameters at some of the last 20 EnKF-based algorithm iterations saved and using either only the cost function, J_{LICA} (left), or only the cost function, J_{RICA} (right), as the simplified model output.

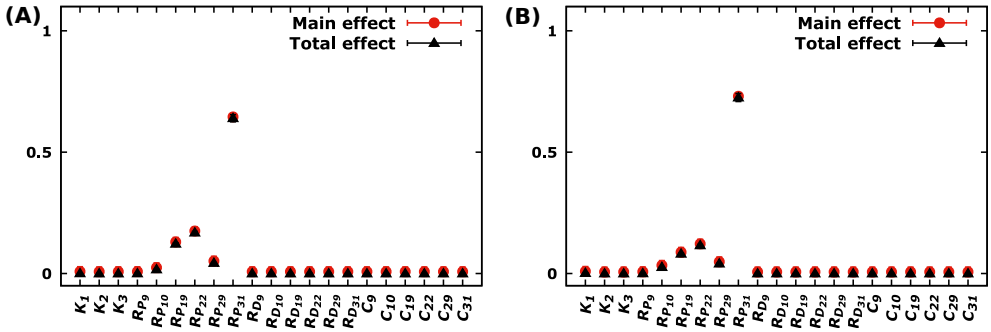

Figure 5.10: Average over the last 20 EnKF-based algorithm iterations of the first order and total indices for the 21 estimated parameters using at each iteration either only the cost function, J_{LICA} (A), or only the cost function, J_{RICA} (B), as the simplified model output.

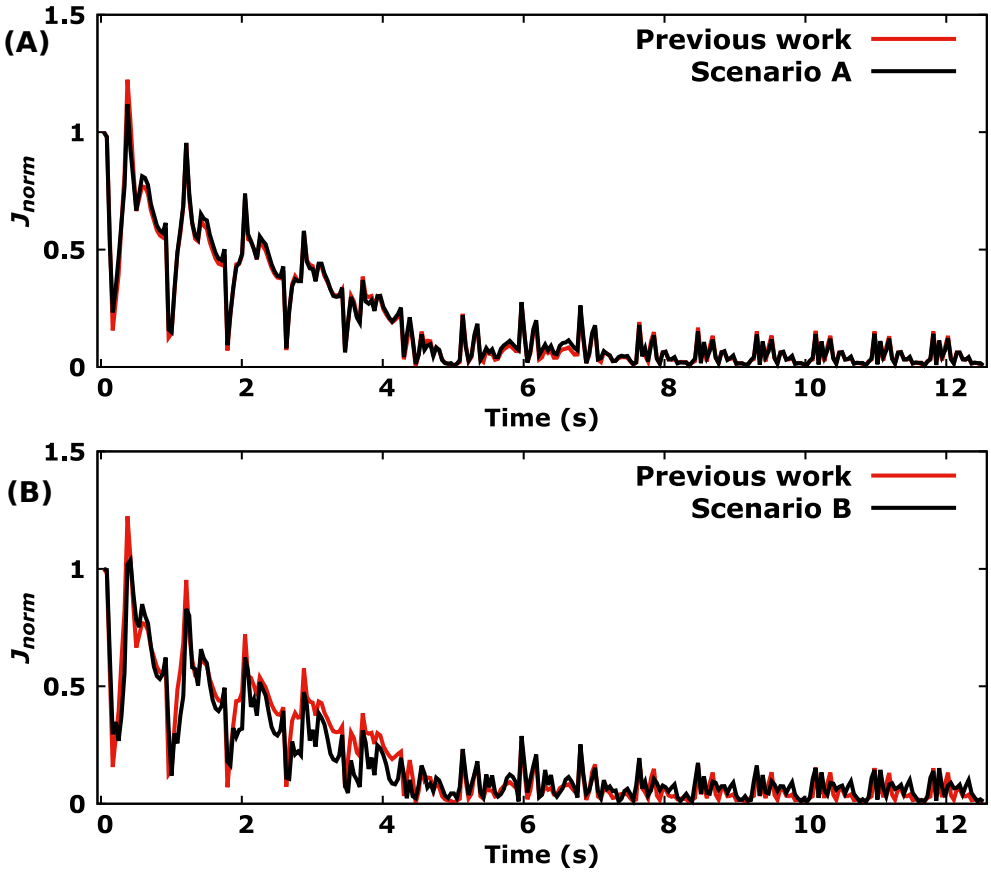

Figure 5.11: Comparison between the normalized cost function, J_{norm}, obtained in [277] (referred to as Previous work) with those obtained in both scenarii considered: (A) for scenario A and (B) for scenario B. The solid line shows the normalized cost function value over time (i.e., the cost function, J, over all the ensemble Kalman filter (EnKF) based algorithm iterations divided by the cost function, J, at the first EnKF-based algorithm iteration).

Acknowledgment

The authors gratefully acknowledge Prof.-Dr. V. Costalat, Dr. E. Le Bars and Dr. J. Deverdun from the Department of Neuroradiology of the *Centre Hospitalier Régional Universitaire de Montpellier, Gui de Chauliac, Montpellier, France* and Dr. J. Siguenza for providing the medical data.

Chapter key points:

- **The most influential parameters in the inversion outcomes involve most of the proximal resistances.**
- **Most of the proximal resistances are responsible for the uncertainty in the patient-specific blood flow rates estimates in the internal carotid arteries.**
- **A stopping criterion based on the convergence of the most influential parameters of the model is proposed for the EnKF algorithm used in the inversion procedure.**

Part III

Machine learning for blood pressure estimation

CHAPTER 6

Augmented patient-specific functional medical imaging by implicit manifold learning

Chapter contents

This chapter has been published as:

Rapadamnaba, R, Nicoud, F, Mohammadi, B. "Augmented patient-specific functional medical imaging by implicit manifold learning" in *International Journal for Numerical Methods in Biomedical Engineering*, 2020; 36:e3325. https://doi.org/10.1002/cnm.3325 [278]

Material from this chapter has been presented as:

Rapadamnaba R., Nicoud F., Mohammadi B. "Estimation non-invasive de la pression dans les artères cérébrales" at 9ème Biennale Française des Mathématiques Appliquées et Industrielles, 13 - 17 mai 2019, Guidel Plages (Morbihan), France.

The numbering of sections, figures and tables in the original manuscript have been altered for this chapter, to be consistent with the thesis chapter numbering. The references have been incorporated into the global references at the end of the thesis.

Part of Section 6.2 of this chapter has been previously presented in Sections 4.3 and 4.4 of chapter 4 and Section 5.2 of chapter 5. The other sections of this chapter present new materials focusing on the use of a machine learning technique together with a convolutional neural network as an alternative to the EnKF-based approach for the patient-specific cerebral blood pressure estimation.

Abstract

This paper uses machine learning to enrich magnetic resonance angiography and magnetic resonance imaging acquisitions. A convolutional neural network is built and trained over a synthetic database linking geometrical parameters and mechanical characteristics of the arteries to blood flow rates and pressures in an arterial network. Once properly trained, the resulting neural network can be used in order to predict blood pressure in cerebral arteries noninvasively in nearly real-time. One challenge here is that not all input variables present in the synthetic database are known from patient-specific medical data. To overcome this challenge, a learning technique, which we refer to as implicit manifold learning, is employed: in this view, the input and output data of the neural network are selected based on their availability from medical measurements rather than being defined from the mechanical description of the arterial system. The results show the potential of the method and that machine learning is an alternative to costly ensemble based inversion involving sophisticated fluid structure models.

Keywords: machine learning, transfer learning, convolutional neural network, hemodynamic problems, noninvasive pressure estimation.

6.1 Introduction

At the junction of optimization, statistics and computer science, machine learning (ML) has gained tremendous interest throughout all scientific disciplines. One can say without exaggeration that it is a revolution in the way solutions of scientific problems are considered. This revolution is very noticeable in biomedical applications with the works of Rahul, Liang et al., Adam et al., Pozin et al., Koeppl et al., Koprowski et al., Cang et al., Luo et al., Cilla et al. [182, 383, 384, 385, 386, 387, 388, 389, 390], and also more recently, with the works of Goceri, Zhang et al., Chen et al., and Nguyen

and Wei [391, 392, 393, 394, 395], to name just a few. According to several researchers, ML will undoubtedly continue to significantly revolutionize biomedical research and global health care in the future [182, 183, 184]. Indeed, fuelled by increasingly powerful computers and available storage capacities, ML methods are able to deal with large, complex and heterogeneous data typically found in biomedical applications [182, 183, 184, 394, 396, 397]. Thus, ML can contribute to identify and engineer features from the data, perform more robust predictions and therefore lead to more accurate diagnostic algorithms and individualize patient treatments [184, 398, 399, 400, 401, 402].

ML-based strategies most often rely on models applied to a carefully constructed dataset so that the training and test data are drawn from the same feature space and the same marginal probability distribution [243]. As pointed out by Ruder [189], these models sorely lacks the ability to generalize to conditions that are different from those encountered during learning and for which the models might be unable to make reliable predictions. This is the case, for instance, when a model learns from simulations and is being applied to deal with actual data. As data from the real world is messy and may contain several new scenarii, many of which were not present during the learning step, the model tuned from numerical data suffers from a significant loss in performance or even breaks down completely when used in practice [189]. In the field of ML, the ability of a model to transfer knowledge to new conditions is generally known as transfer learning (TL) and this is one of the key features used in this paper.

Many examples exist in the ML literature where TL has been applied successfuly: text categorization, text and Web-document classification [403, 404, 405, 406], learning in real-time strategy game [242], text mining [405, 407], natural language processing [408], WiFi-based Indoor Localization [409], reinforcement learning [410, 411, 412, 413, 414], to cite a few. In this paper, we apply TL, in particular learning from simulations, to the prediction of blood pressure in a cardiovascular system.

Learning from simulations is one particular application of TL commonly used in ML when one needs to deal with data, which is either rare, incomplete, expensive or simply dangerous to obtain [189]. In all these situations, analytic models or numerical simulations can fill the gap. In previous works, the use of an algorithm based on ensemble Kalman filter (EnKF) technique coupled to a 0D compartement network has allowed to noninvasively estimate patient-specific blood pressure in cerebral arteries [12, 277]. However, this procedure, which requires inverse problem solving, is too long, expensive and difficult to integrate in a medical acquisition device. The present work aims at remedying that with focus on the following three key messages:

- Use supervised ML together with a Convolutional Neural Network (CNN) — a type of ML algorithms used with great success in miscellaneous image recognition applications and prediction tasks, such as handwriting recognition and face detection [246, 261, 264, 415] — as an alternative to the ensemble based inversion procedure by EnKF used in our previous works for cerebral blood pressure estimation. In particular, on the basis of magnetic resonance angiography and magnetic resonance imaging (MRA & MRI) acquisitions provided by physicians and using a learning database generated with the same numerical procedure as that used in the inversion procedure, the aim is to recover the EnKF outcomes previously achieved by Lal et al. [12] through the solution of an inverse problem. This allows to augment/enrich already available information in MRA & MRI images.

- Use "implicit learning on a manifold" to deal with situations where part of the information in the input variables are missing while observations are available on some output variables. This is why the term "implicit" is introduced. The term "manifold" is used to reflect the fact that among all the variables contained in the database, only a subset of them will be used for the learning. This is like working on a vector subspace and consists of reorganizing the database into *structured* variables in order to take into account unavailable patient-specific data.

Therefore, "implicit learning on a manifold" should not be seen as a particular Artificial Intelligence (AI) algorithm. In fact, it should be regarded as just a multi-output convolutional neural network. Thus, any multi-output regression method (e.g. linear models, multi-output random forests or other forms of neural nets with autoencoding) could have been used instead of this multi-output convolutional neural network implementation.

- Apply the outcome of the learning from simulations to patient-specific data and we use the term "transfer" in that sense. The network and its coefficients obtained during the learning from simulations are kept unchanged when applied on patient-specific data.

The remainder of the paper is organized as follows. First, an overview of the methodology followed in this work is provided in Section 6.2. Section 6.3 details the procedure for building the synthetic database. Section 6.4 presents the state of the art in ML and discusses some mathematical aspects of the CNN approach used in this study. Section 6.5 illustrates the effectiveness of the method by considering a very simple problem while Section 6.6 shows its application for the estimation of blood pressure in cerebral arteries. Finally, Section 6.7 discusses some of the limitations of the present study and Section 6.8 provides concluding remarks.

6.2 Preliminaries

6.2.1 General methodology

Figure 6.1 presents the general flowchart for this work with the following main steps:

1. **Hemodynamic and morphological data extraction**: first, from magnetic resonance angiography and magnetic resonance imaging acquisitions and segmentation of a 3D time of flight magnetic resonance angiography (3D-TOF-MRA) — `dicom` files provided by the Department of Neuroradiology

of the *Centre Hospitalier Régional Universitaire de Montpellier (CHRU), Montpellier, France*, — blood flow rates in ascending aorta (AA), right and left internal carotid arteries (R-ICA and L-ICA) have been extracted using the `GTFlow` sofware[1] together with the morphological data relevant to geometric measurements and the morphology of some arteries through the use of the RadiANT DICOM Viewer software[2].

2. **Patient-specific arterial network construction**: on the basis of these images, a patient-specific arterial network of 33 arteries as shown in Figure 6.2 consisting of the aorta, vertebral, carotid and brachial arteries together with the complete circle of Willis has been then constructed. For more details, the reader is referred to previous works [12, 277].

3. **Machine learning**

   a) **Database generation**: the next step has consisted in generating synthetic data by performing a series of forward simulations of the blood flow model ($\mathcal{M}$) described in detail below in Section 6.2.2. The outcome of these simulations relates the *model input parameters* (namely, the blood flow rates in AA, the geometrical (radius r, thickness h, and length l) and mechanical characteristics (Young moduli) of the 33 arteries as well as the Windkessel three-element parameters) to the *model output parameters* (namely, the blood flow and pressure in the 33 arteries.)

   b) **The use of CNN as ML algorithm**: subsequently, a multi-output CNN-based approach was introduced in order to learn on *structured* variables from the database, the relationship between CNN input and output parameters, and then, apply the acquired knowledge to patient-specific data for blood pressure estimation.

[1] Available on `https://www.gyrotools.com/gt/index.php/products/gtflow`.
[2] See `http://www.radiantviewer.com/`.

4. **Validation**: Finally, to test the accuracy of the method, pressure estimated using CNN was compared with that previously predicted in Lal et al. [12] by solving an inverse problem using EnKF.

Compared with the method used in previous works [12, 277], the novelty in the general framework proposed here is the use of ML method instead of EnKF-based parameter estimation algorithm.

6.2.2 The blood flow model

Mathematical modeling is a powerful tool for a better understanding of the cardiovascular system in an inexpensive and noninvasive way. There are several cardiovascular models (0D, 1D, 3D) available in the literature, each coming with its advantages and drawbacks. In this work, we adopted a 0D formulation, also known as lumped-parameter model, due to its simplicity and effectiveness in describing first order features of the human cardiovascular system.

In general, the mathematical formulation of the lumped-parameter models is derived by decomposing the cardiovascular system into single arterial segments or compartments, each segment being represented with a single electrical circuit, consisting of resistance R, inductance L, and compliance C, as shown in Figure 6.3. In this way, the full arterial network is built by connecting single electrical compartments together.

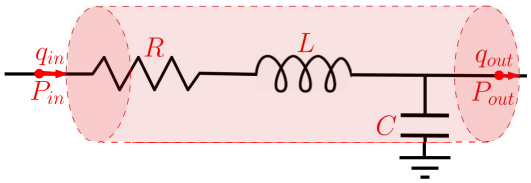

Figure 6.3: A single compartment circuit illustration.

Applying the laws of conservation of mass and momentum to a single compartment filled with an incompressible Newtonian fluid, the following linear first order ordinary differential equations can be obtained:

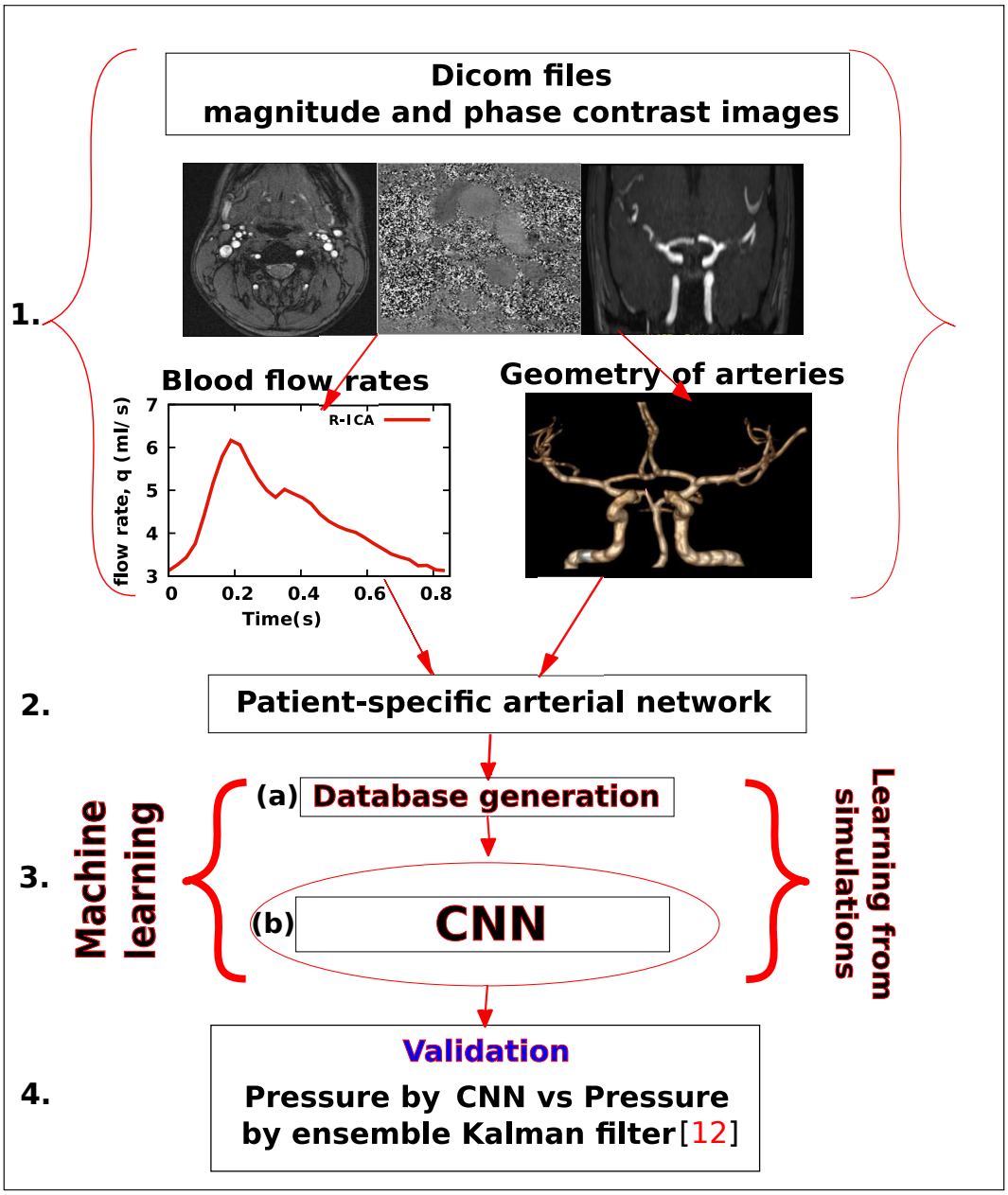

Figure 6.1: The general framework of the proposed methodology. The numbers on the left side refer to the four steps of the general methodology described above.

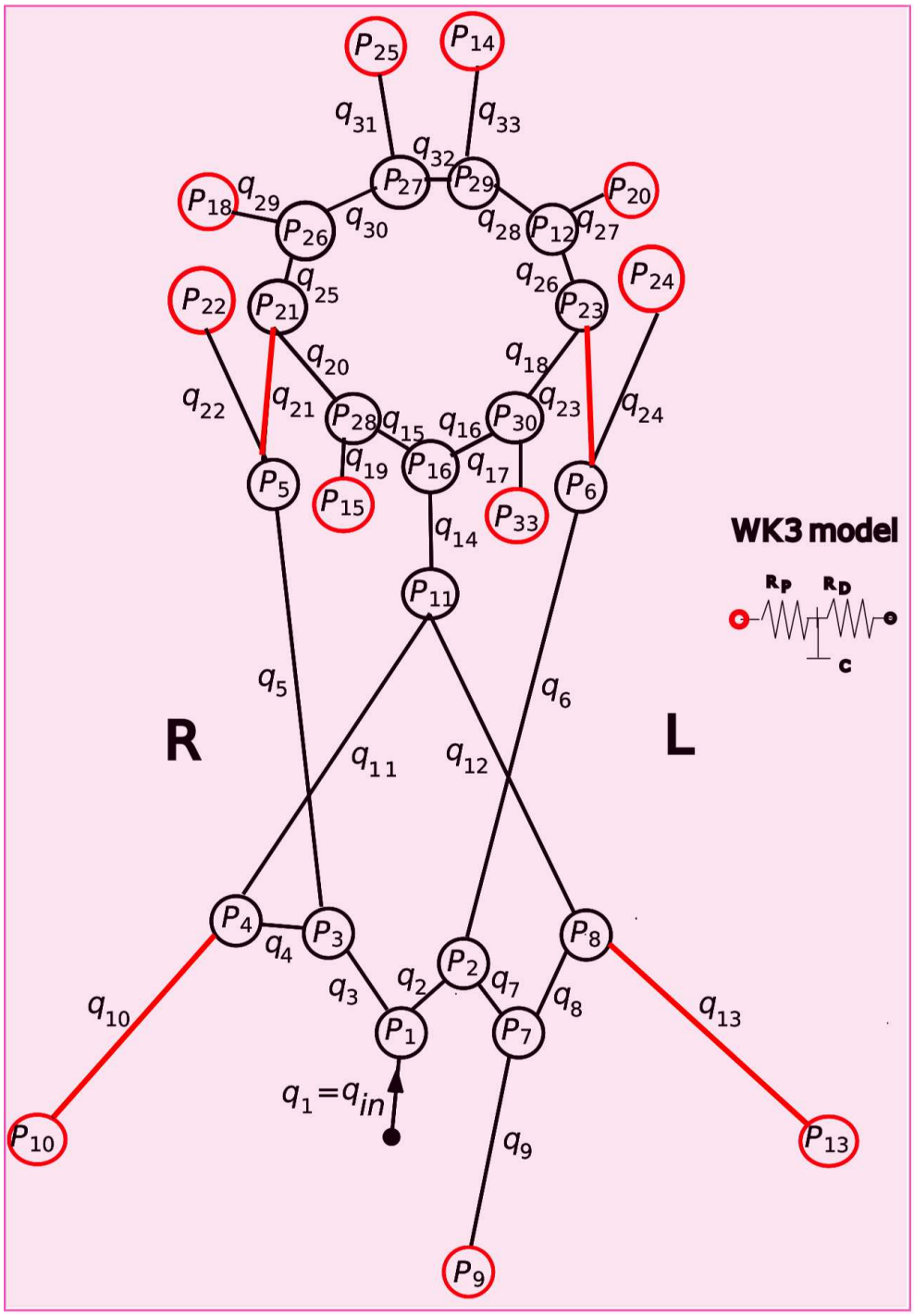

Figure 6.2: The 0D arterial network model [12] of the patient-specific upper body arteries and the circle of Willis. The arterial segments in red and the red circles represent the four arteries of interest examined in the rest of paper and the WK3 model, respectively.

$$\begin{cases} C\dot{P}_{out} = & q_{in} - q_{out} \\ L\dot{q}_{in} = & P_{in} - P_{out} - Rq_{in}, \end{cases} \tag{6.1}$$

where (P_{in}, q_{in}) and (P_{out}, q_{out}) refer to the pressure and flow rate at the inlet and outlet of the vascular segment, respectively. For arteries with a radius $r < 0.2\,\text{cm}$, the inertial effect is ignored [317] and the flow rate is given by $q_{in} = (P_{in} - P_{out})/R$. Let r, h, l, μ, ρ, and E be the radius, the wall thickness, the length, the blood viscosity, the blood density and the Young modulus of the arterial segment, respectively. The three parameters R, L, C are obtained from physical principles (momentum and mass conservation) applied to the arterial sector and read [86]:

- the viscous flow resistance, R, is derived from the Poiseuille's equation, which describes the relation between pressure drop, ΔP, and the steady blood flow, q, through a uniform and rigid blood vessel: $R = 8\mu l/\pi r^4$;
- the inductance of the vessel, L, which represents the acceleration and deceleration of the inertial blood at each heart beat, is equal to $L = \rho l/\pi r^2$;
- the vessel compliance, C, which represents the ability of a vessel wall to expand and contract with changes in pressure, is equal to $C = 3\pi r^3 l/2Eh$.

In order to close the arterial network and accommodate the cumulative effects of vessels beyond the terminal segments, the three-element Windkessel model (WK3-lumped parameter model) [286, 318] is prescribed at the outlet of each terminal segment. As a reminder, the equation relating the instantaneous blood pressure and the flow rate reads as follows:

$$\dot{p}(t) + \frac{p}{R_D C} = R_P\,\dot{q}(t) + \frac{qR_T}{R_D C}, \tag{6.2}$$

where p, q, C, R_P, and R_D refer to the instantaneous pressure at the inlet of the WK3 model, the instantaneous flow rate, the compliance, the proximal resistance, and the distal resistance of the vascular beds, respectively, and R_T refers to the total peripheral

resistance meaning that $R_T = R_P + R_D$. For simplicity, the blood flow model — let us refer it to as $(\mathcal{M})$ — can be formally expressed as follows :

$$(\mathcal{M}) : \begin{cases} \dot{Y}(t) = & F_X(Y(t)) \\ Y(0) = & Y_0, \end{cases} \tag{6.3}$$

where $Y(t) = (P(t), q(t))^\top$, Y_0 is the initial condition, X represents the geometrical parameters of the problem (i.e r, h, l), the inlet flow rate from AA, the Young moduli, the Windkessel WK3 boundary conditions [12] at the same time, and F_X is a linear or nonlinear function defined by the blood flow model describing the relation between $Y(t)$ and its time derivatives $\dot{Y}(t)$. Here and for the rest of the paper, the notation 'function with suscript X', for examle F_X, means that the function F depends on the parameter X. This is a generic formalism for any 0D arterial network and therefore includes the 0D formulation of the cardiovascular network investigated in the current study (Figure 6.2). This network is taken from Lal et al. [12] because both studies focus on the same patient.

6.3 In silico database generation

In this section, we describe the procedure for generating synthetic data. This data consists of geometrical parameters (the radius, r, the wall thickness, h, and the length, l), physical parameters (the Young moduli, E), blood flow rates, q, blood pressures, P for the 33 elements of the arterial model (Figure 6.2) on top of the eleven WK3 models used as boundary conditions (the proximal resistances, R_P, the distal resistances, R_D, and the compliances, C).

For a forward simulation using the blood flow model $(\mathcal{M})$ presented in Section 6.2.2, this data is interrelated in the sense that the variables r, h, l, E, R_P, R_D, C of the network associated with the variable q_{in}, the inlet flow rate (the blood flow rate in AA), represent the input parameters of the model, which we refer to as *model input parameters*, and the variables q and P, the output parameters, which we refer to as

model output parameters. As a reminder, the Young moduli E are given by $Eh = r(k_1\, e^{k_2 r} + k_3)$ [321] where k_1, k_2, k_3 are model constants and the inlet flow rate q_{in} is defined using two components α and τ as follows [12, 277]:

$$q_{in}(t) = \begin{cases} \alpha \sin(\pi t/\tau) & \text{if } t < \tau, \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

where $\alpha = 302$ ml/s and $\tau = 0.3$ s are nominal values taken from Lal et al. [12]. These will be randomly perturbated below following a Gaussian distribution with means equal to these nominal values and standard deviations σ, which will be specified later, to produce samples of q_{in}. For more details, the reader is referred to Lal et al. [12].

To generate the database, blood rheological parameters were fixed as: the blood density $\rho = 1050\,\mathrm{kg{\cdot}m^{-3}}$ and the blood viscosity $\mu = 0.004\,\mathrm{Pa{\cdot}s}$, and the same WK3 boundary conditions for left and right pairs of terminal compartments were assumed. This means that the terminal compartments such as the segment connecting P_{27} and P_{25} and the one linking P_{29} and P_{14} in Figure 6.2, for instance, are assigned with the same WK3 boundary conditions. A series of forward simulations of the blood flow model ($\mathcal{M}$) is then performed using a series of perturbed *model input parameters* ($\tilde{q}_{in}$, $\tilde{r}$, $\tilde{h}$, $\tilde{l}$, $\tilde{E}$, $\tilde{R}_P$, $\tilde{R}_D$, $\tilde{C}$). The series of outcomes obtained from these simulations ($P = \tilde{P}_1, \cdots, \tilde{P}_{33}$, $q = \tilde{q}_1, \cdots, \tilde{q}_{33}$) associated with the series of perturbed *model input parameters* constitutes the database. This is illustrated on Figure 6.4. Note that, while not expressly stated, all the parameters in the database representing blood flow rates and pressures are time-dependent. It is also worth noting that the parameters in the database can be subdivided into two categories: the class of known parameters (easy to measure) and the class of unknown parameters (difficult to measure). Table 6.1 provides a general overview of these different parameters listed by category.

To obtain the series of perturbed *model input parameters* for each forward simulation of the blood flow model ($\mathcal{M}$), we randomly perturbate each *model input parameter* following a Gaussian distribution with a mean equals to the nominal value of each input parameter taken from Lal et al. [12], and a standard deviation σ of 5% of this

nominal value. The standard deviation σ of 5% of each nominal value is chosen in order to obtain perturbated *model input parameters* tending to realistic values encountered in the literature. E being given by the relation $Eh = r(k_1\, e^{k_2 r} + k_3)$, $\tilde{E}$ is obtained by perturbating the parameters appearing in the relation giving E meaning that $\tilde{E}\tilde{h} = \tilde{r}(\tilde{k}_1\, e^{\tilde{k}_2 \tilde{r}} + \tilde{k}_3)$, where each of the 3 constants $\tilde{k}_1$, $\tilde{k}_2$, and $\tilde{k}_3$ is also obtained from a Gaussian distribution with a mean equals to the nominal value of each constant k_1, k_2, and k_3 taken from Lal et al. [12] and a standard deviation of 5% of this nominal value.

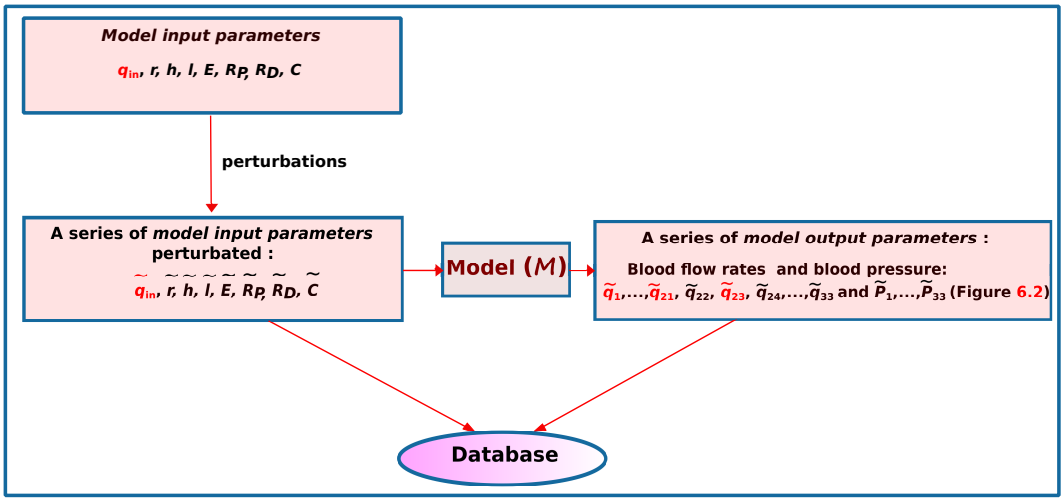

Figure 6.4: Illustration of the database generation procedure. In the case studied here, the parameters in red are considered as known parameters (easy-to-measure parameters) while the remainder is considered to be unknown.

Table 6.1: The two categories of parameters in the database: the known parameters (easy-to-measure parameters) and the unknown parameters (not-easy-to-measure parameters).

Known parameters	**Unknown parameters**
blood flow rates in AA, $\tilde{q}_{in} = \tilde{q}_1$	geometrical parameters of the 33 arteries, $\tilde{r}_1, \cdots, \tilde{r}_{33}$, $\tilde{h}_1, \cdots, \tilde{h}_{33}$, $\tilde{l}_1, \cdots, \tilde{l}_{33}$
blood flow rates in R-ICA, $\tilde{q}_{21}$	Young moduli of the 33 arteries, $\tilde{E}_1, \cdots, \tilde{E}_{33}$
blood flow rates in L-ICA, $\tilde{q}_{23}$	blood pressures in the 33 arteries, $\tilde{P}_1, \cdots, \tilde{P}_{33}$
—	blood flow rates in 30 arteries, $\tilde{q}_2, \cdots, \tilde{q}_{20}, \tilde{q}_{22}, \tilde{q}_{24}, \cdots, \tilde{q}_{33}$

6.4 The state of the art in machine learning

A tremendous number of ML algorithms have been developed, which all share the same goal: to provide sufficient flexibility to minimize training error but, at the same time, allow generalization to new data sets, all in a computationally efficient way [182]. Among these algorithms, the most commonly used ones are Linear Regression, Logistic Regression, Linear Discriminant Analysis, Decision Tree, Support Vector Machine (SVM), Naive Bayes, k-Nearest Neighbors (kNN), Learning Vector Quantization (LVQ), K-Means, Bagging, Random Forest, Dimensionality Reduction Algorithms, Convolutional Neural Networks (CNN), Gradient Boosting algorithms and AdaBoost [187, 416, 417, 418]. These algorithms tasks mainly include regression, classification, predictive modeling, clustering, association or survival analysis, link mining, and dimensionality reduction [187, 190, 419]. The details of these different algorithms are beyond the scope of this paper. However, CNN is discussed below because of its ability to manage complicated relationships between input and output data that are not easily captured by manual measurement [183], its relevance to biomedical research, its frequent use within the literature [184] and its use within the current work.

6.4.1 The implicit CNN : notation and database destination

The first step in our approach is to receive from a user raw data made of a set of scenarii linking input and output variables (let us refer to them as input data and output data, respectively). In the sequel, as we are interested in biomedical applications, the considered database will consist of a set of N scenarii linking n input real variables (these are typically blood flow rates values at temporal discretization points) to m output real variables (these also typically correspond to blood pressure values at temporal discretization points): $\{X \in \mathbb{R}^n \longrightarrow Y \in \mathbb{R}^m\}$ where $X = (X_1, ..., X_n)$ and $Y = (Y_1, ..., Y_m)$. Hence, the required data is as follows:

$$X^1 = (X_1^1, ..., X_n^1) \longrightarrow Y^1 = (Y_1^1, ..., Y_m^1),$$

$$\vdots$$

$$X^N = (X_1^N, ..., X_n^N) \longrightarrow Y^N = (Y_1^N, ..., Y_m^N).$$

In the applications we are interested in, data is often rare, expensive to obtain or confidential. The aim, therefore, is not to handle very large amount of data, but rather see the potential of learning of CNNs based on small amount of data in opposition to big data situations. The destination of the data is rather classical. The database is splitted in three subsets to be used to train, validate and *a posteriori* test the network. In this splitting, the major part of data is used for learning. A second subset is used to make sure to avoid over-fitting during learning (this is called validation). The third subset is for *a posteriori* test of the quality of the network. This kind of splitting is very classical in statistical data treatment and widely described in dedicated manuscripts [420].

6.4.2 The implicit CNN algorithm construction

Classical CNN algorithms are usually built to mimic and replicate the investigated physical model process. This means that in the classical CNN, the input and output data of the neural network are exactly the same as those used in the physical model considered. This is illustrated in Figure 6.5 in the case examined here, where the input data and the output data of the network are directly derived from the series of perturbated *model input parameters* and from the series of *model output parameters*, respectively. Note that except the blood flow rates $\tilde{q}_{in}$ and ($\tilde{q}_{21}$, $\tilde{q}_{23}$) in the input data and the output data, respectively, the other parameters constituting the input/output data in the classical CNN are unknown (not easy to measure). This prevents us from performing the *a posteriori* test of the quality of the network, hence the idea of changing the structure of the classical CNN algorithm so that all the known parameters in the database can be gathered as consisting of the network input data.

The CNN algorithm proposed here, which we refer to as implicit CNN, is not usual. One of its particularities is that its input data is neither usual nor classical. Indeed, in contrast to the classical approach, the input and output data of the network in the

implicit CNN algorithm does not mimic those from the physical model. They are rather reorganized in *structured* variables. The example shown in Figure 6.6 demonstrates such organizational change in the implicit CNN : the input data consists of one subset of the series of perturbated *model input parameters*, $\tilde{q}_{in}$, and one subset of the series of *model output parameters* (blood flow rates $\tilde{q}_1, \cdots, \tilde{q}_{33}$), and the output data consists of the remaining part of the series of *model output parameters* (blood pressure $\tilde{P}_1, \cdots, \tilde{P}_{33}$). However, in practice, we are interested in predicting the blood pressure in a specific artery and not in all at once. Thus, instead of using all available information as shown in Figure 6.6, a simpler construction where the input and output data of the implicit CNN will consist of $\tilde{q}_{in}$ and the series of blood flow rates in the artery of interest on the one hand, and on the other hand, the series of blood pressure in the same artery (see Figures 6.10 – 6.13 below) takes over from it. This unusual construction is motivated by the available parameters from the patient's MRA & MRI images. Indeed, only the parameters available (known parameters) within the patient-specific data (see Table 6.1), namely the parameters extracted from MRA & MRI images and not requiring the resolution of an inverse problem — the blood flow rates in AA, R-ICA and L-ICA, $\tilde{q}_{in}$, $\tilde{q}_{21}$, and $\tilde{q}_{23}$ (see Figure 6.2), respectively are used as input data of the network. The idea behind this construction is to obtain a network capable of linking only the easy-to-measure parameters (known parameters in *model input parameters*) from the patient's MRA & MRI images to the *model output parameters*. This implies that the information in the unknown parameters from *model input parameters* which are not used for building the network is implicitly contained in the *model output parameters* used for the network construction and can therefore be indirectly recovered. This means that exploiting only known parameters from *model input parameters* as input data of the network allows to access to hidden information in the unknown parameters not used. In addition to allowing the *a posteriori* test of the quality of the network, this construction avoids the problem of unknown parameter estimation through inverse problem solving before performing the convolutional neural network construction.

Another particularity of our implicit CNN concerns the definition of the network. In the classical CNN techniques, the number of layers and the number of hidden variables in each layer must be *a priori* defined by the user before the parameters of the network are found through error backpropagation. This is known to require lots of know-how by the user. We avoid this *a priori* definition introducing an incremental procedure for the definition of the network structure. We proceed in a systematic way combining an incremental construction of the network with the use of the information available through activation functions. This means that different network depths are tested with increasing number of layers. For each of these networks, a maximum number of variables per layer is prescribed. In our case, we consider a maximum allowance of 200 hidden variables for all layers. Eventually, after this optimization procedure, only the variables which are found to sufficiently contribute to the network outcome are retained. These correspond to those with the activation function beyond a given user-defined threshold. The network we retain is the one with the best fitting capability over the learning database while avoiding overfitting. This optimization of the network structure is interesting for both the optimization and inference steps. For the former, once a given variable is found not being of sufficient importance, it is removed from the optimization set. This means that search activities along this variable are abandoned hence reducing the size of the optimization space. For the latter, a lighter network will obviously permit more effective inference steps in term of computational effort. Figure 6.7 illustrates this procedure for the simple example given in Section 6.5. It shows the evolution of the mean average error over the test database for increasing network depth and the evolution of the cumulative number of the variables of the network. The network, which shows the best compromise between accuracy and complexity is found to be of depth 20: adding extra layers and variables do not significantly improve the quality of the results. The results presented below in Figure 6.9 corresponds to the outcome of this network. Note also that in our implicit CNN, 1D convolution operations with Gaussian kernels are used in convolution layers

as we handle time series. The supports of the convolutions are also automatically identified in the same way as the optimal number of layers of the network.

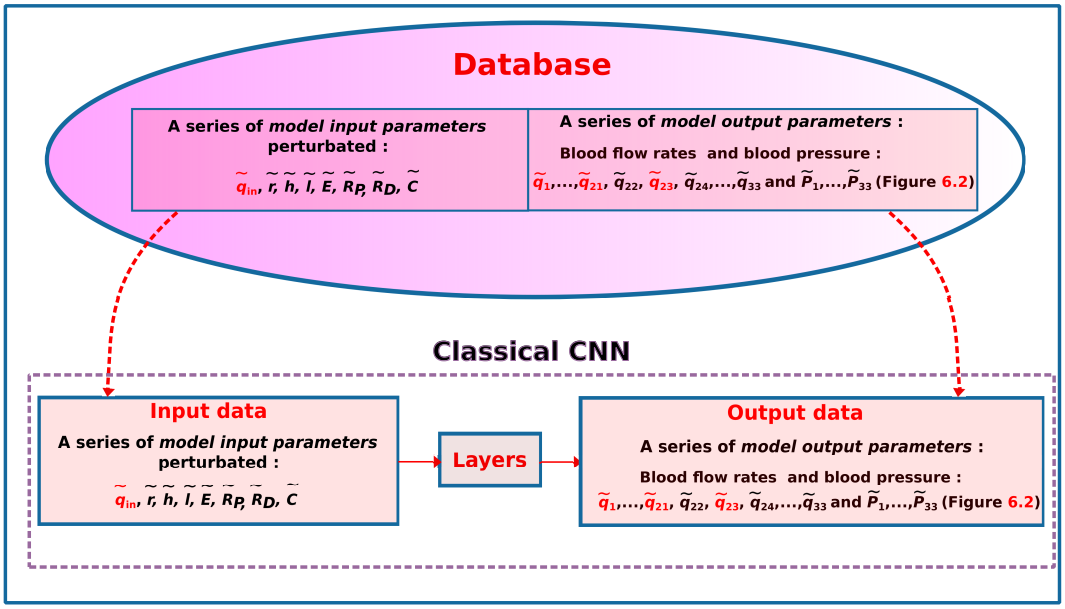

Figure 6.5: The classical CNN construction. The network input data consists of a series of *model input parameters* perturbated $\tilde{q}_{in}, \tilde{r}, \tilde{h}, \tilde{l}, \tilde{E}, \tilde{R}_P, \tilde{R}_D, \tilde{C}$, and the network output data is composed of the corresponding series of *model output parameters* ($\tilde{q}_1, \cdots, \tilde{q}_{33}$ and $\tilde{P}_1, \cdots, \tilde{P}_{33}$ in Figure 6.2). The known parameters are in red and the unknown ones are in black.

6.5 Illustration on a simple example: CNN and solution of a forward problem

Let us illustrate the proposed approach on the solution of an ordinary linear differential equation:

$$\begin{cases} \dot{p}(t) &= f_X(t), \\ \dot{q}(t) &= g_X(t), \\ p(0) &= p_0, \\ q(0) &= q_0, \end{cases} \tag{6.5}$$

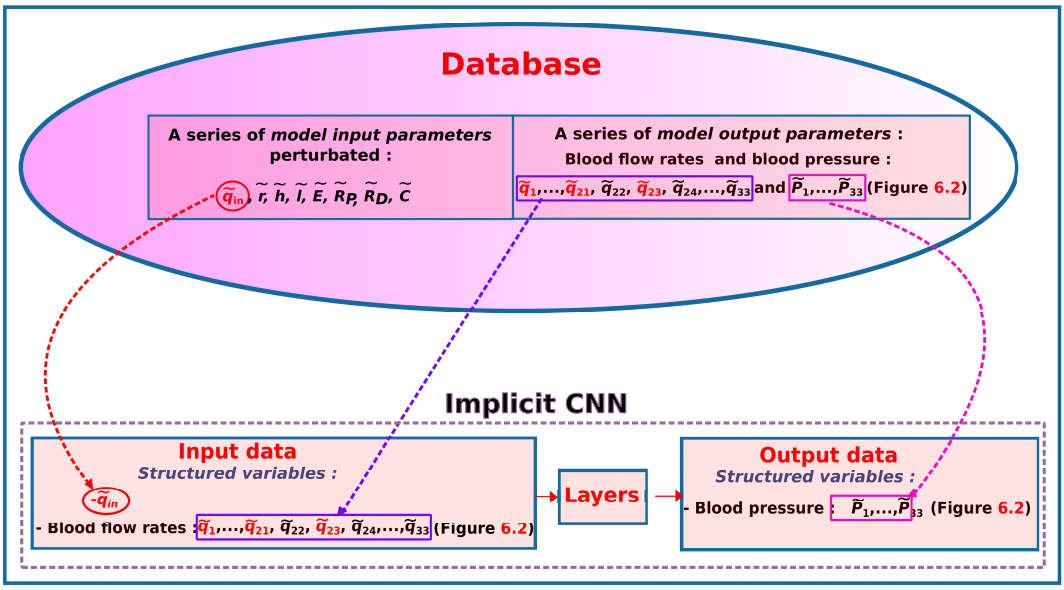

Figure 6.6: The implicit CNN construction. The network input data consists of a series of *structured* variables: the series of known parameters from the series of *model input parameters* perturbated, namely the perturbated inlet flow rate, $\tilde{q}_{in}$, and the corresponding series of blood flow rates from the series of *model output parameters* ($\tilde{q}_1, \cdots, \tilde{q}_{33}$ in Figure 6.2), and the network output data is composed of the corresponding series of blood pressure from the series of *model output parameters* ($\tilde{P}_1, \cdots, \tilde{P}_{33}$ in Figure 6.2). The known parameters are in red and the unknown ones are in black.

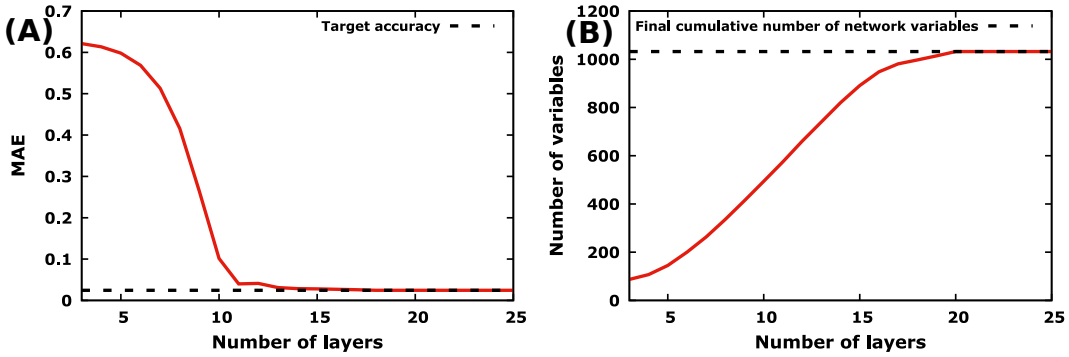

Figure 6.7: Incremental implicit network construction for the simple example given in Section 6.5. A: Mean average errors versus number of layers and B: total number of network cumulative variables versus number of layers.

with $X = (X_1, X_2)$, $f_X(t) = X_1 \cos(X_1 t + X_2)$, $g_X(t) = -X_1 \sin(X_1 t + X_2)$, $p_0 = \sin(X_2)$, and $q_0 = \cos(X_2)$.

Setting $Y = (p, q)^\top$, $H_X(t) = (f_X(t), g_X(t))^\top$, $Y(0) = (p(0), q(0))^\top$ and $Y_0 = (p_0, q_0)^\top$, this ordinary linear differential equation can be restated as:

$$\begin{cases} \dot{Y}(t) &= H_X(t), \\ Y(0) &= Y_0. \end{cases} \tag{6.6}$$

The solution Y_* of the latter differential equation is trivial and we have $Y_*(t) = (p_*(t), q_*(t))^\top = (\sin(X_1 t + X_2), \cos(X_1 t + X_2))^\top$. In order to apply CNN-based technique to this simple problem, a database containing samples of $X \in \mathbb{R}^n$ and $Y \in \mathbb{R}^m$, is digitally generated.

For the classical CNN, we have $n = 2$, the size of input data X. Moreover, the solution Y_* is represented by Y (output data), a vector of size $m = 86$ corresponding to the number of points in a discrete representation of the solution Y_* (43 for each of the two components of Y). In this test case, the database contains $N = 200$ samples of Y obtained for 200 values of X. In the applications we are interested in, $N = 200$ different scenarii is already quite large as data is often not that easy to obtain. This choice of 200 samples is not coerced. This is arbitrary and we could, of course, have chosen any other reasonable number of samples.

For the implicit CNN, data is reconstructed as follows: the input data X' is built from a subset of X and a subset of Y. For instance, one can consider $X' = (X_1, q)$. The output data Y' is made of the remaining part of Y, p. Thus, the size of input data is now $n = 44$ (1 for X_1 and 43 for q), and that of output data is $m = 43$. Note that X_2 belongs neither to the input nor the output but its effect on the output is present, although not explicitly written.

The results obtained using the classical CNN and the implicit CNN for this simple problem are illustrated in Figures 6.8 and 6.9, respectively. In these Figures, the upper parts (Figures 6.8A and 6.9A) show the networks determined to properly represent the actual "physical" system described in Eq. 6.6 : a network consisting of 30 locally fully

connected layers with up to 155 hidden variables for the classical CNN, and a network of 20 locally fully connected layers with up to 84 hidden variables for the implicit one. By locally fully connected we mean that the variables of two successive layers are connected. In the remaining part of these Figures (Figures 6.8B, C and 6.9B), the desired solution is referred to as Target, the prediction of the desired solution referred to as CNN, and data used for learning referred to as Learning data. Comparing the Target and CNN data, the results are very good with a relative error of less than 3 % and mean absolute errors (MAE) of 0.05 and of 0.024 for the classical and implicit CNN, respectively. Note that the implicit CNN is able to recover the partial output p without using the explicit knowledge of the X_2 input. This is a big advantage compared to the classical CNN formulation, especially in any case where X_2 is difficult to measure.

6.6 Application of the implicit CNN to 0D blood flow model

6.6.1 A test case with synthetic data

We are interested in the blood flow model ($\mathcal{M}$) presented in Section 6.2.2 and we would like to build a database allowing to estimate the blood pressure in some arteries of interest. In practice, for the hemodynamic applications we are interested in, some parameters are measurable and others not. In the case investigated here, blood flow rates in AA, R-ICA and L-ICA (q_{in}, q_{21} and q_{23}, respectively), and systolic and diastolic blood pressure in left and right brachial arteries (SBP and DBP in L.BRA and R.BRA, the maximum and the minimum value of P_{10} and P_{13}), were measured. This information led to choose both ICAs and both brachial arteries as the arteries of interest. To estimate the blood pressure in these four arteries of interest, we first generate the database following the procedure described in Section 6.3. From this database, four different CNNs, one for each artery of interest and each based on $N = 100$ samples, were built following the implicit CNN structure for a single artery of interest described in Section 6.4.2.

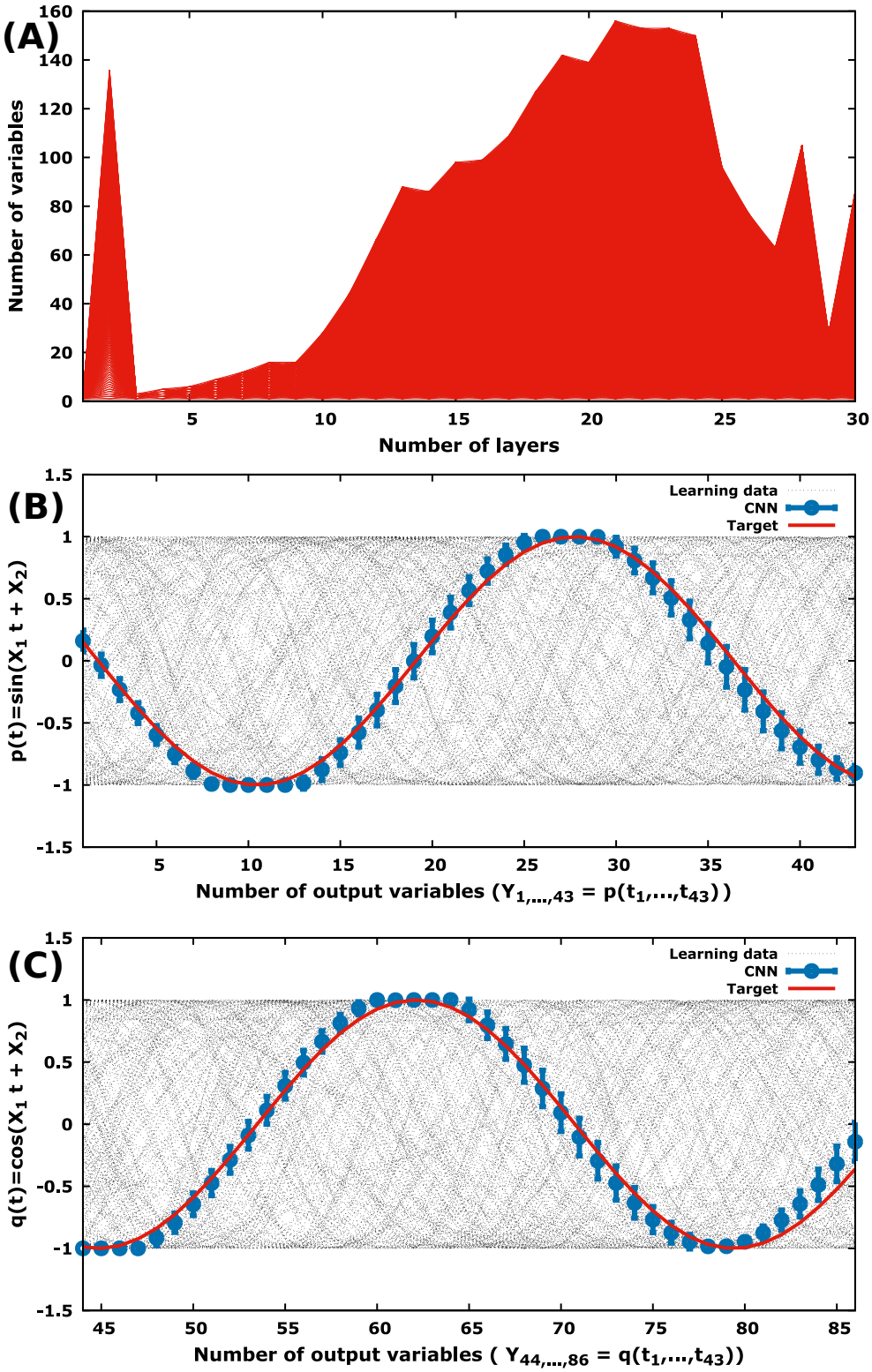

Figure 6.8: Learning the solution of an ordinary differential equation using a database of 200 available solutions with the classical CNN. A: the network determined for this problem consists of 28 locally fully connected hidden layers with up to 155 hidden variables, and B,C: the results of the target solution prediction using the classical CNN.

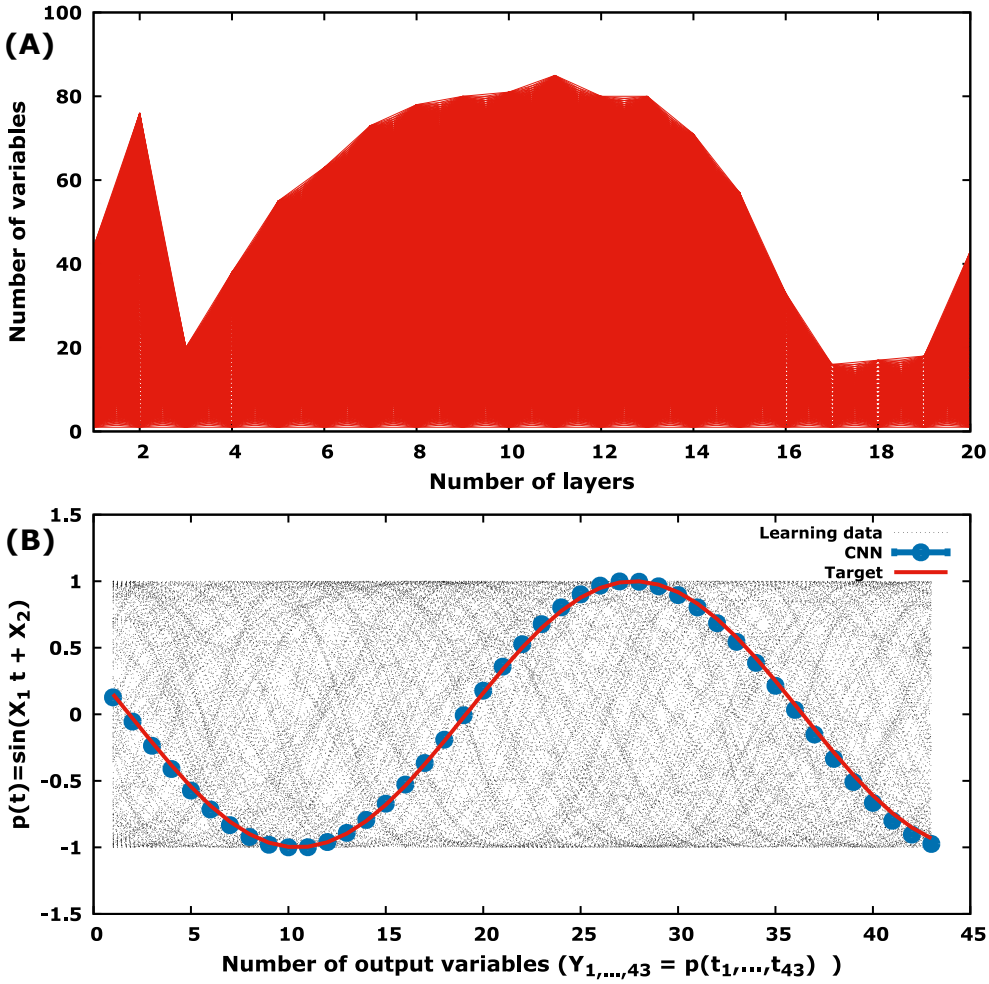

Figure 6.9: Learning the solution of an ordinary differential equation using a database of 200 available solutions with the implicit CNN. A: the network determined for this problem consists of 18 locally fully connected hidden layers with up to 84 hidden variables and B: the results of the target solution prediction using the implicit CNN.

To build each CNN, a relationship is created between the input data of size 34 (2 for the two components $\tilde{\alpha}$ and $\tilde{\tau}$ of blood flow rates in AA, $\tilde{q}_{in}$, and 32 for the number of points in a discrete representation of the blood flow rates in the artery of interest) and the corresponding output data of size 32 representing the number of points in a discrete representation of blood pressure in the artery of interest. To be more precise, Figures 6.10,6.11,6.12, and 6.13, referred to as CNN R-ICA, CNN L-ICA, CNN R.BRA, and CNN L.BRA, respectively, show the implicit CNN structure for each of the four arteries of interest (Figures 6.10 and 6.11 for ICAs, and Figures 6.12 and 6.13 for brachial arteries). Thus, for each artery of interest, the input data size is $n = 34$, the output data size is $m = 32$, and the number of samples is $N = 100$.

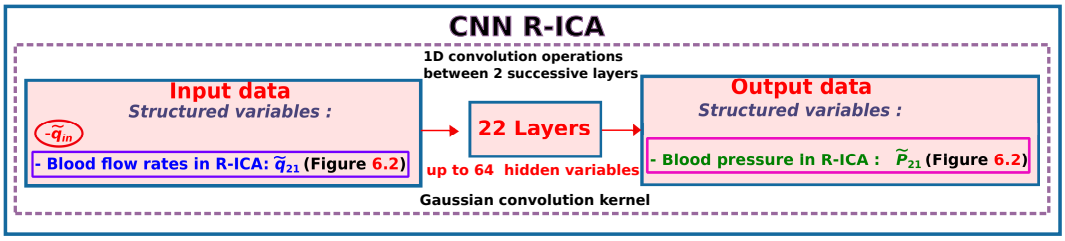

Figure 6.10: The implicit CNN for R-ICA. The network input data consists of a series of perturbated inlet flow rates parameters $\tilde{q}_{in}$ and of the corresponding series of blood flow rates in R-ICA ($\tilde{q}_{21}$ in Figure 6.2) and the network output data is composed of the corresponding series of blood pressure in R-ICA ($\tilde{P}_{21}$ in Figure 6.2).

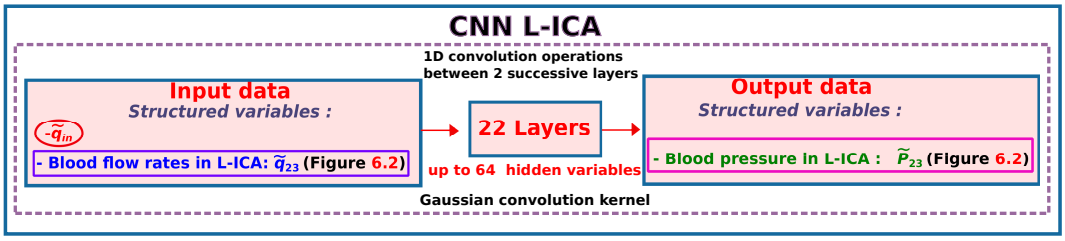

Figure 6.11: The implicit CNN for L-ICA. The network input data consists of a series of perturbated inlet flow rates parameters $\tilde{q}_{in}$ and of the corresponding series of blood flow rates in L-ICA ($\tilde{q}_{23}$ in Figure 6.2) and the network output data is composed of the corresponding series of blood pressure in L-ICA ($\tilde{P}_{23}$ in Figure 6.2).

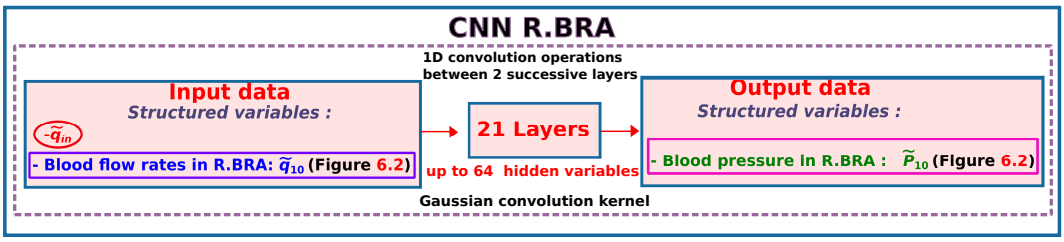

Figure 6.12: The implicit CNN for R.BRA. The network input data consists of a series of perturbated inlet flow rates parameters $\tilde{q}_{in}$ and of the corresponding series of blood flow rates in R.BRA ($\tilde{q}_{10}$ in Figure 6.2) and the network output data is composed of the corresponding series of blood pressure in R.BRA ($\tilde{P}_{10}$ in Figure 6.2).

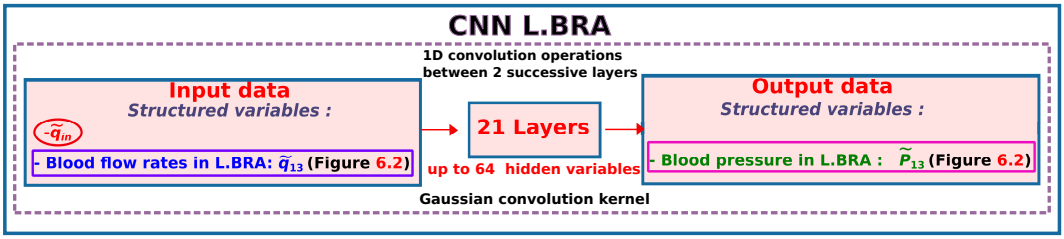

Figure 6.13: The implicit CNN for L.BRA. The network input data consists of a series of perturbated inlet flow rates parameters $\tilde{q}_{in}$ and of the corresponding series of blood flow rates in L.BRA ($\tilde{q}_{13}$ in Figure 6.2) and the network output data is composed of the corresponding series of blood pressure in L.BRA ($\tilde{P}_{13}$ in Figure 6.2).

Each of the four implicit CNNs built as in Figures 6.10-6.13 is trained using 99 samples whose the output data is referred to as Learning data in Figure 6.14. The last sample whose the output data is referred to as Target (the desired outcome) is used as testing sample. The results obtained are shown in Figure 6.14 which illustrates the networks used for R-ICA (Figure 6.14A) and for the left brachial artery (Figure 6.14B). Prediction of the desired blood pressure is referred to as CNN (Figure 6.14C, E for ICAs, and Figure 6.14D, F for brachial arteries). It is worth pointing out that for the two networks illustrated here, about twenty locally fully connected layers with up to 64 hidden variables have been effective in providing good results, and that using CNN with the same number of layers and hidden variables, similar results were observed for L-ICA (Figure 6.14E) and R.BRA (Figure 6.14F). From Figure 6.14, the predicted blood pressure recovers quite well the target blood pressure. Indeed, in all cases, the relative

error between CNN prediction and the target values is less than 4%. Moreover, for the ICAs, the relative error for systolic and diastolic blood pressure is less than 1%. As for the brachial arteries, they have a relative error of less than 2% as well for systolic blood pressure as diastolic one.

To assess the relevance of these results and evaluate the skill and the performance of the four implicit CNNs on new data, a traditional approach is to undertake k-fold cross-validation. This approach consists first in splitting the $N = 100$ samples of the database into k roughly equal-sized subsamples [421]. Then, one of the k subsamples is retained as testing set, and the remaining subsamples are used as learning data. This process is then repeated k times with each of the k subsample being used for learning. Thus, all the subsamples are used for both learning and testing and each subsample is used for testing exactly once. For the choice of k, even if the choice k=10 is very common in the field of applied ML, and is recommend if one is struggling to choose a value for the dataset, there is no formal rule [396, 422, 423]. However, as pointed out by Kuhn, Gareth, and many others before them, in practice, one perfoms k-fold cross-validation using k=5 or k=10 [396, 423, 424, 425]. In fact, Both values of k have been shown empirically to yield test error rate estimates that suffer neither from excessively high bias nor from very high variance [396]. Given these considerations, k-fold cross-validation technique is performed here using k=5, 10, and 100. The case k=100, also known as leave-one-out cross-validation (LOOCV), is also taken into account since this case generalizes the results presented above where 99 samples make up the training samples and only one sample is used as testing sample. Still, for each k-fold cross-validation run, the MAE obtained are averaged to produce a single estimation, which we refer to as the k-fold mean absolute errors (k-MAE). Table 6.2 shows typical k-MAE values obtained for blood pressure estimation using the implicit CNN in R-ICA and L.BRA and the corresponding relative errors with respect to the Target diastolic and systolic blood pressures (DBP and SBP). Just for the record, these Target values (DBP and SBP, in mmHg) were 110.67 / 161.49 for R-ICA, and 117.59 / 172.73 for L.BRA.

The results in Table 6.2 shows that the k-MAE values decrease with the increase in the size of the training dataset meaning that more learning data leads to fewer errors in the prediction. This is consistent with the results in ML literature. Moreover, the relative errors with respect to the Target values for the case k=100 (less than 2%) seems to be comparable with the results obtained above without cross-validation using 99 samples for learning and a single sample for testing. These results being quite encouraging, the next step will focus on applying the implicit CNN to patient-specific data.

Table 6.2: The k-fold mean absolute errors, k-MAE for k=5, 10, 100 and the corresponding relative errors with respect to the Target systolic and diastolic blood pressures (SBP and DBP) values in R-ICA and L.BRA. This shows typical skill scores of the implicit CNN for in-vitro pressure prediction in R-ICA and L.BRA.

Arteries of interest	**R-ICA**, mmHg			**L.BRA**, mmHg		
	k-**MAE**	**Relative errors**		k-**MAE**	**Relative errors**	
		DBP	**SBP**		**DBP**	**SBP**
k=**5**	**7.36**	**6.65%**	**4.56%**	**8.58**	**7.3%**	**4.97%**
k=**10**	**5.47**	**4.94%**	**3.38%**	**4.81**	**4.09%**	**2.78%**
k=**100**	**2.03**	**1.83%**	**1.26%**	**2.17**	**1.84%**	**1.26%**

6.6.2 A test case using patient-specific PC-MRA & MRI-based data

The patient-specific data used in the current study has been provided by the Department of Neuroradiology at the *Centre Hospitalier Régional Universitaire de Montpellier (CHRU), Montpellier, France.* Arterial systolic and diastolic blood pressures at rest of a healthy volunteer were measured before and after image acquisition using a brachial automatic sphygmomanometer (Maglife, Schiller Medical). The systolic and diastolic values were 125 mmHg and 72 mmHg in the right brachial artery, and 115 mmHg and 72 mmHg in the left brachial artery. 2D phase-contrast imaging was performed on a Siemens 3T Skyra MR Scanner. The ascending aorta and the internal carotid arteries (right and left) 2D phase-contrast images were considered for the patient-specific blood flow rates extraction. For more details, the reader is

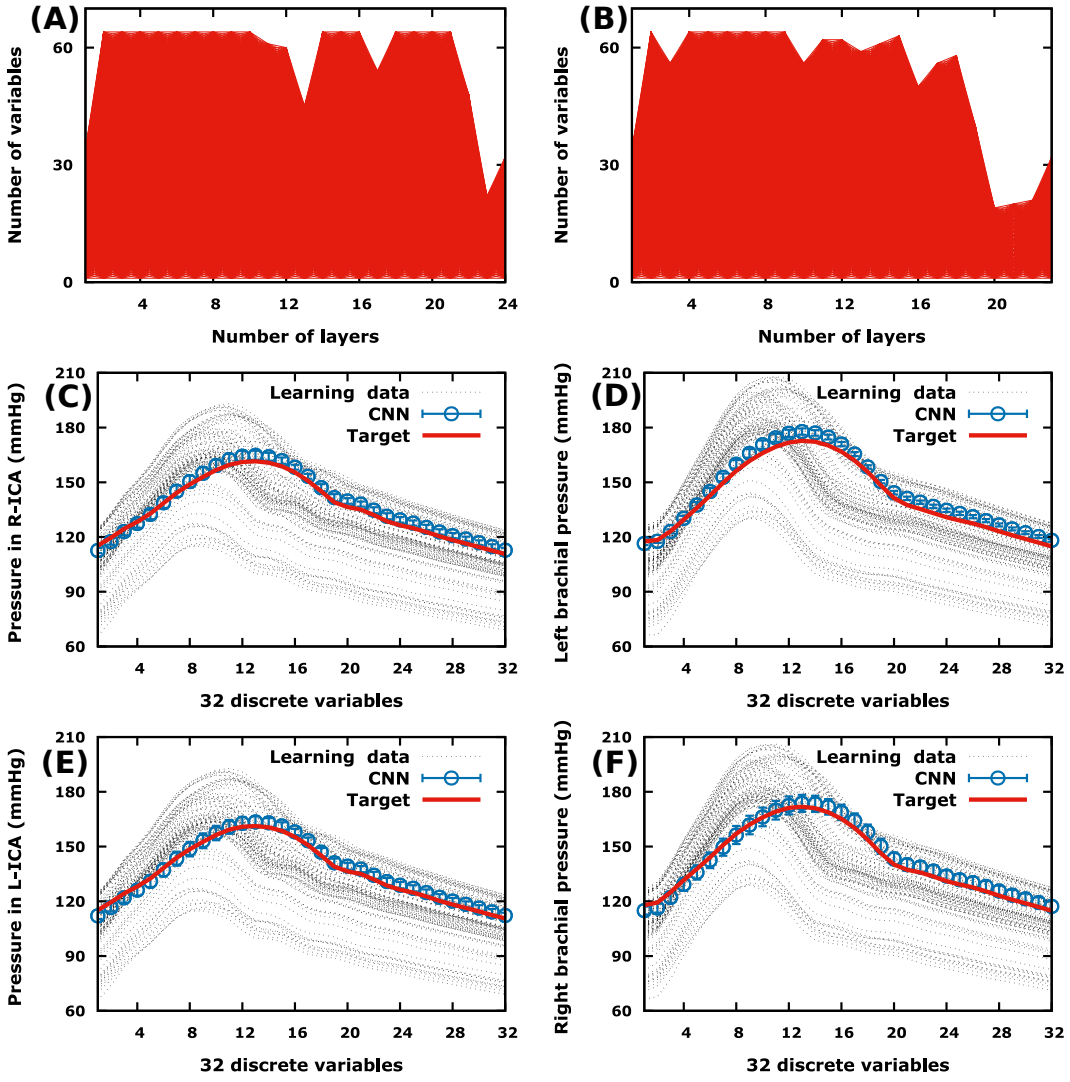

Figure 6.14: Typical networks determined once for all for estimating blood pressure in R-ICA (A) and in left brachial artery (B), estimated blood pressure in the four arteries of interest (in R-ICA (C), in left brachial artery (D), in L-ICA (E), and in right brachial artery (F)) using four different implicit CNNs ($n = 34, m = 32$, and $N = 100$). For R-ICA and L-ICA, the network determined has 22 hidden layers with up to 64 hidden variables, and for the brachial arteries, there are 21 hidden layers.

referred to Lal et al. [12].

For each artery of interest, the implicit CNN built for the synthetic case (Section 6.6.1) is reused with patient-specific data. More precisely, the implicit CNN algorithm applies the acquired knowledge during learning in the synthetic case namely, the relationship between input and output data in each CNN previously built in the synthetic case (the network and its coefficients obtained during the learning from simulations are kept unchanged) to patient-specific data. In fact, the testing samples in the synthetic case are replaced with patient-specific data from the different arteries of interest. Thus, each testing sample consists now of the 2 variables (α and τ) modeling the patient-specific inlet flow rate, and the 32 points in a discrete representation of the patient-specific blood flow rates in the artery of interest. As a reminder, R-ICA, L-ICA, R.BRA, and L.BRA are here the arteries of interest. For R-ICA and L-ICA, the implicit CNNs obtained in the synthetic case are applied to the patient-specific blood flow extracted from MRA & MRI in the ICAs, which is used as input data. The resulting output, the blood pressure estimated via CNN in the ICAs is then compared to the desired outcome, the blood pressure previously estimated by Lal et al. [12] in the same arteries but through the solution of an inverse problem using EnKF.

In the same way, for R.BRA and L.BRA, the implicit CNNs built in the synthetic case are applied to patient-specific blood flow but this time not to those extracted from MRA & MRI (measurements not available in these arteries) but to those previously predicted in the brachial arteries by Lal et al. [12] with the blood flow model ($\mathcal{M}$) using EnKF. Again, in this case, the resulting outcome, the blood pressure predicted in the brachial arteries using CNN, is compared to that previoulsy predicted by Lal et al. [12] in the same arteries through the solution of an inverse problem using EnKF. Comparison with the brachial pressure measurements is also performed to test the accuracy of the method used.

Figure 6.15 shows a comparative overview between typical results of blood pressure

predicted via CNN (referred to as CNN) and typical results of blood pressure previously predicted *via* EnKF (referred to as EnKF). It is observed that secondary peaks in the blood pressure waveforms are well reproduced with CNN. The same trend was observed for L-ICA and R.BRA. In all cases, for the ICAs, the CNN simulated waveforms have a relative error of less than 3% in the diastolic and systolic blood pressure. Tables 6.3 and 6.4 summarize the cardiac cycle systolic and diastolic blood pressure estimated using CNN and that simulated using EnKF, in the four arteries of interest (Table 6.3, for the ICAs and Table 6.4, for the brachial arteries). From these Tables, all systolic and diastolic blood pressure predicted using EnKF and CNN results differ by less than 7%. In addition, in order to give an idea of the prediction error made by CNN with respect to available blood pressure measurements, the predicted systolic and diastolic pressure values in the brachial arteries are compared with patient-specific blood pressure measurements. As reported in Table 6.4, the relative error obtained using CNN is less than 3% for diastolic pressure and in the two brachial arteries, less than 6% for systolic pressure in R.BRA but up to 16 % for systolic pressure in L.BRA. One explanation for this larger error is that the input/output testing sample for this case is already from previous numerical results — the results perfomed by Lal et al. [12] — and not from patient-specific data extracted from MRA & MRI. Moreover, note that there is a significant difference between the systolic blood pressure value previously predicted by Lal et al. [12] using EnKF in L.BRA (129.32 mmHg) — what needs to be predicted here — and the systolic pressure measurement value (115 mmHg). Apart from this extreme case, the agreement between the predicted and measured pressure values is rather good. This shows that CNN can effectively be a good alternative to the previous method proposed by Lal et al. [12] for blood pressure estimation through the solution of an inverse problem using EnKF. This is interesting in terms of cost and computation time. Indeed, learning takes place in 2 minutes on a 3GHz CPU and a single inference in 0.05 second. The learning time includes the identification of the suitable network structure. This inference time should be compared to the 7 hours needed by the ensemble based

inversion using EnKF. In other words, if the unit time is the computation time for pressure estimation by CNN, EnKF requires 240,000 times more computation time. A reason for this extra cost in EnkF is that the EnKF-based technique needs to retrieve first the mechanical properties of arteries before providing the arterial pressures while machine learning does not need any identification of the mechanical properties of the arteries in order to retrieve the arterial pressures.

Table 6.3: Validation: comparison of the cardiac cycle systolic blood pressure (SBP) and diastolic blood pressure (DBP) estimated using CNN with the model simulated values using EnKF in the ICAs. The percentage difference between the results obtained using CNN and those based on EnKF is stated in bold in parenthesis.

ICA	Systolic blood pressure (SBP), mmHg		Diastolic blood pressure (DBP), mmHg	
	CNN	EnKF	CNN	EnKF
Right	114.80 (±0.00) **(0.46%)**	114.28	69.23 (±0.04) **(2.32%)**	67.656
Left	114.54 (±0.00) **(0.46%)**	114.02	69.30 (±0.04) **(2.49%)**	67.617

Table 6.4: Validation: comparison of the cardiac cycle systolic blood pressure (SBP) and diastolic blood pressure (DBP) estimated using CNN with the model simulated values using EnKF in the brachial arteries. The predicted pressure values using CNN are also compared to patient-specific brachial pressure measurements. The percentage difference between the results obtained by CNN and those based on EnKF, and the percentage difference between the results obtained by CNN and the blood pressure measurements are stated in bold in parenthesis.

BRA	Systolic blood pressure (SBP), mmHg			Diastolic blood pressure (DBP), mmHg		
	CNN	measurement	EnKF	CNN	measurement	EnKF
Right	131.92 (±0.59) **(3.55%)**	125 **(5.53%)**	127.4	73.54 (±0.18) **(6.57%)**	72 **(2.14%)**	69
Left	133.71 (±0.60) **(3.39%)**	115 **(16.27%)**	129.32	73.21 (±0.17) **(6.63%)**	72 **(1.68%)**	68.66

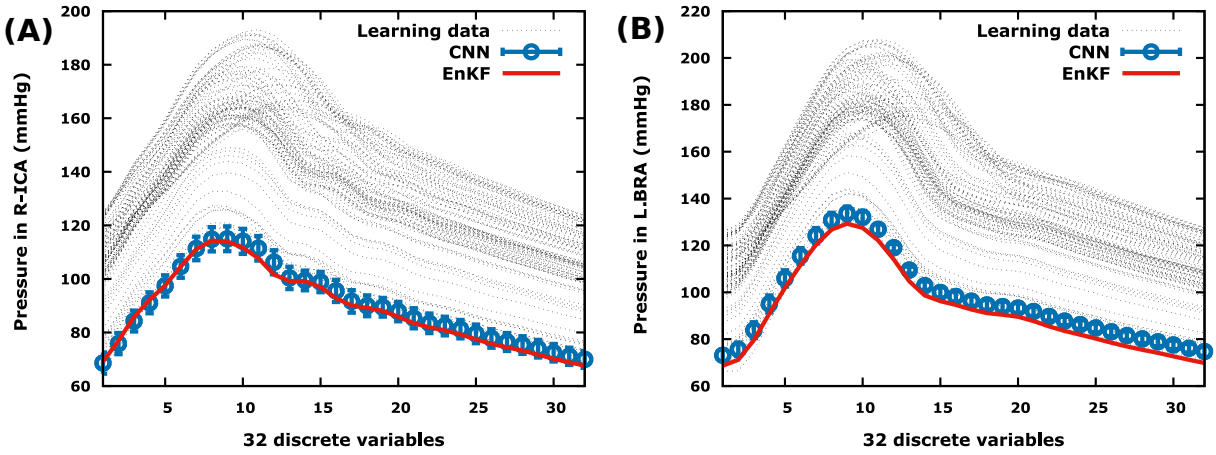

Figure 6.15: Prediction of the solution of a cardiovascular problem (patient-specific blood pressure) in R-ICA (A) and in L.BRA (B) using the network built during training in the synthetic case, in Section 6.6.1. The 99 available simulated solutions (blood pressure) are referred to as Learning data, the predicted pressure as CNN, and the pressure previously estimated using EnKF [12], as EnKF.

6.7 Discussion and limitations

The objective of the present investigation is to use machine learning in order to propose an alternative to an ensemble based inversion procedure (EnKF) for cerebral blood pressure estimation. The presented implicit manifold learning technique is found to be much faster than EnKF yet providing very similar results in terms of blood pressure signals. The *in vitro*/*in vivo* results show good agreement between the target blood pressures and the outcome of the machine learning procedure. In particular, the previously predicted SBP and DBP in the four arteries of interest using EnKF are well recovered with machine learning procedure.

One of the limitations could be the use of a simple cardiovascular model comprising of 0D components and the modeling of inflow q_{in} using Equation 6.4. Indeed, several studies have pointed out the shortcomings of lumped models and of the use of this approximation of inflow and have proposed more complex/accurate models to better capture the characteristics of blood flow. So we are aware that more sophisticated physical models can be considered. However, to be able to make a fair comparison, we seek a supervised learning approach encapsulating the same physics as the one used in

our previous studies [12, 277] with the objective to get the cerebral blood pressure in nearly real time. Again, this is why the same cardiovascular model and the arterial network have been considered as with the EnKF procedure. Of course this procedure should receive more testing in a clinical environment.

One particularity of this work is that unlike most learning techniques, the implicit learning proposed here is operational with small data. This should be seen as a constraint and not something we fully control. Data might come from other teams and we do not always have control of it. Moreover, the only thing we can be sure of, regardless of the AI algorithm one uses, is that the quality of a database decreases with its size because it simply becomes more difficult to guarantee the quality of the data when the size of the sample is becoming large. Data quality issue can also be related to possible inconsistency between scenarios (e.g. low and high fidelity prediction of a same situation). This is a classical problem with supervised learning. Therefore, we think that working with small database is desirable when this is possible. Furthermore, being able to produce meaningful results with small database will permit to operate in cases where the learning is carried out not with synthetic data but with real experiments data, which is obviously more difficult to multiply.

Another particularity of this work is that, in order to make learning and inference more efficient, learning has been performed on a manifold, using only a subset of the parameters in the database. One natural question is then how this kind of restriction can affect the results. To address this question, ablation analysis is performed in the real data case. An ablation study refers to removing some feature of a model in order to see how that affects the outcomes. As a reminder, the implicit CNN relates only the inflow q_{in} and blood flow rate in any artery of interest to blood pressure in the same artery. But, the available information from MRA & MRI images are inflow q_{in} and blood flow rates in R-ICA & L-ICA. So, when the implicit CNN is performed for predicting blood pressure in R-ICA, the blood flow rate in L-ICA is systematically ablated and vice-versa. Now all flow rates information available from MRA & MRI images (namely,

inflow q_{in} and blood flow rates in R-ICA & L-ICA) and blood flow rates previously estimated by Lal et al. [12] in R.BRA and L.BRA associated with inflow q_{in} are used as CNN input data, respectively. This permits to see how the networks output is affected by the addition of this extra information. The results presented in Figure 6.16 show that the construction is quite robust with respect to such ablation.

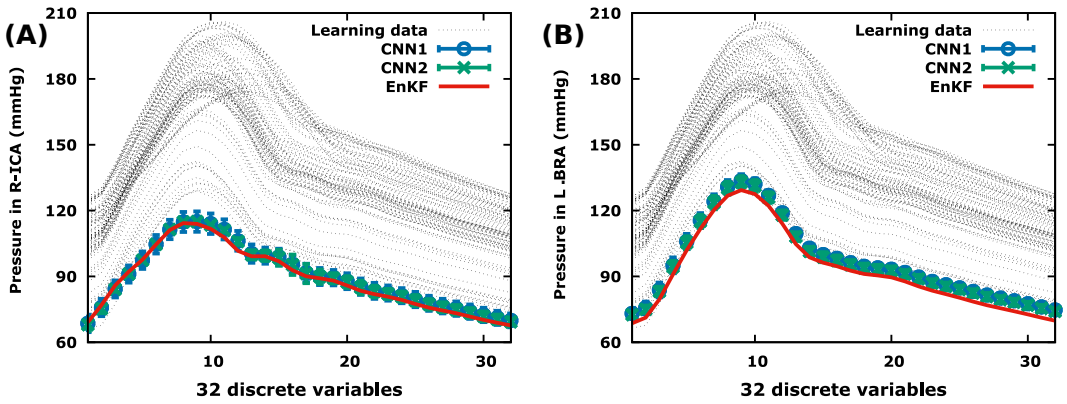

Figure 6.16: Ablation analysis for the patient-specific case: CNN1 and CNN2 results are practically the same. CNN1 represents the predicted blood pressure using parts of the input features (inflow q_{in} and blood flow rate in a particular artery of interest, in R-ICA (A) and in L.BRA (B) here). This corresponds exactly to what is referred to CNN in Figure 6.15. CNN2 is the predicted blood pressure using together inflow q_{in} and blood flow rates in R-ICA & L-ICA as input data (A) or using at the same time inflow q_{in} and blood flow rates previously estimated by Lal et al. [12] in R.BRA and L.BRA as input data (B).

6.8 Concluding remarks

The main challenge in this paper has been to use machine learning for blood pressure prediction in cerebral arteries. We aim our proposal to be faster and more practical than when the pressure is recovered through the solution of an inverse problem using ensemble methods such as EnKF, yet as accurate. Hence, machine learning has been applied to the solution of cardiovascular problems together with a transfer learning strategy using in-silico databases. Indeed, as patient data is rare, a mathematical model has been used to simulate the relationship between blood flow rates and blood pressure in an human arterial network. The machine learning tool has been applied to patient-specific blood

pressure estimation to provide a direct link in any couple of arterial blood flow rates and pressure in the upper body region. The proposed method demonstrated compelling results on synthetic and patient-specific data. The results suggest that the proposed ML strategy is well-suited for the considered cardiovascular problem providing real-time estimation of the pressure in patient-specific arteries, which is missing in standard MRA & MRI acquisition. In comparison with the EnKF-based pressure estimator, the suggested ML-based estimator provides a substantial gain as regards calculation time. In that sense, we can consider inference by this machine learning approach to provide pressure distribution in real time augmenting the MRA-MRI information on the fly.

Acknowledgment

The authors greatly thank Prof.-Dr. V. Costalat, Dr. E. Le Bars and Dr. J. Deverdun from the Department of Neuroradiology of the *Centre Hospitalier Régional Universitaire de Montpellier, Gui de Chauliac, Montpellier, France* and Dr. J. Siguenza for providing the medical data and for helpful discussions. The blood flow model ($\mathcal{M}$) used for the generation of the database has been solved using the implicit numerical integration solver `DVODE` [319, 320] available on `http://www.radford.edu/~thompson/vodef90web/`.

Chapter key points:

- **ML used together with a CNN is an efficient alternative to the inversion procedure for patient-specific blood pressure estimation in cerebral arteries in nearly real-time.**
- **The ML technique proposed here, referred to as implicit CNN, allows to deal with situations where part of the information in the input variables are missing while observations are available on some output variables.**
- **The methodology used provides practical insights for efficiently and automatically finding the number of variables and layers of NNs.**
- **In comparison with the inversion procedure based estimator, the suggested ML-based estimator provides a substantial gain as regards calculation time.**

Conclusion

CHAPTER 7

Main results, discussions and perspectives

Chapter contents

In this last chapter, the main findings and conclusions drawn from this thesis are recalled. Then, the physical modeling choices and their consequences are discussed. Finally, possible directions for future research are proposed.

7.1 Global conclusion

This thesis was an opportunity, on the one hand, to evaluate the robustness and accuracy of the EnKF-based inversion procedure previously implemented by Lal et al. [12] for patient-specific blood pressure estimation in cerebral arteries and, on the other hand, to propose an alternative methodology capable of providing as accurate and faster patient-specific blood pressure estimation in cerebral arteries as in the inversion procedure.

The investigation of the inversion procedure accuracy and robustness through uncertainty and sensitivity analyses included first the effect of using non-deterministic inlet flow rates, the impact of a change in boundary conditions from Windkessel WK3 model to WK1 model, and the effect of the assumed symmetry in the terminations boundary conditions on the blood flow model results (estimated blood flow rates and pressures).

The blood flow model outcomes related to sensitivity analysis with respect to the inlet flow rate did not show any reduction or amplification of the assumed uncertainty introduced in the inlet flow rate. On the contrary, the predicted blood pressure and elastic parameter were found to be of the same order of magnitude as the assumed uncertainty on the inlet flow. This meant that the inversion procedure seemed to work linearly with respect to the uncertainty propagation and therefore, gave some indication for the level of uncertainty one could tolerate because of manual acquisition steps.

As for the impact of a change in the boundary conditions from WK3 to WK1 on the inversion procedure, the blood flow model outcomes were found to be sensitive to such a shift. Indeed, this analysis showed that the WK1 boundary conditions were clearly insufficient to describe the underlying physics of the model. Indeed, despite the inversion's success in recovering the target flow rate at R-ICA and L-ICA segments with WK1 boundary conditions, the convergence of the EnKF algorithm was not fully achieved and physiological blood pressure values could not be obtained (brachial pressure values were miscalculated). Therefore, WK3 boundary conditions appeared necessary for a correct retrieval of the target signals and for good physiological results.

Regarding sensitivity of the procedure with respect to the assumed symmetry in the terminations boundary conditions, the results revealed that available observations were not enough discriminating so that it was not obvious for the patient-specific case studied here to identify possible nonsymmetry through the inversion procedure.

To conclude this first sensitivity investigation, a low-complexity backward uncertainty quantification construction using a local linear relationship between the blood flow rate and the hemodynamic parameters was introduced to provide an alternative technique to EnKF approach for estimation of covariance matrix of the arterial network hemodynamic parameters. This construction, which can be used with any deterministic inversion algorithm and which is specifically useful when ensemble approach could be unsuccessful, was favourably compared to the EnKF outcomes.

The other investigation of the inversion procedure accuracy and robustness through uncertainty and sensitivity analysis was oriented more towards the identification of the patient-specific model input parameters exerting the most influence on the model results and the incorporation of stopping criterion in the inversion algorithm by means of the Sobol' sensitivity indices. It was found through this analysis that most of proximal resistances were the most important parameters in the blood flow model. Indeed, the estimated blood flow rates in the internal carotid arteries appeared most sensitive with respect to five of six proximal resistances used in the blood model. This led to set a stopping criterion, based on the convergence of these five key parameters identified, for the EnKF-based algorithm used in the inversion procedure, avoiding unnecessary oversolving and providing faster results from the inversion procedure.

Another dimension of this work was the alternative approach, based on implicit CNNs, for estimating blood pressure more quickly but also accurately in the same patient and with the same clinical data as in the inversion procedure. This alternative technique was investigated through successful applications in several idealised cases (analytical functions and synthetic data), and in the case of the same patient-specific as in the inversion procedure (real clinical data from a healthy volunteer). Solutions of

simple ordinary differential equations were very well estimated and time series synthetic and patient-specific blood pressures in internal carotid and brachial arteries were predicted using only time series synthetic and patient-specific blood flow rate values as input parameters of the CNNs. The predicted blood pressure using the implicit CNNs was then compared with the predicted blood pressure obtained using the inversion procedure. This comparison demonstrated that the CNN-based approach was able to provide results consistent with the inversion procedure in terms of brachial and internal carotid arteries pressure estimation for the same patient-specific data. Indeed, in all cases most of blood pressure waveform features were well recovered: the secondary peaks (dicrotic notch) were well recovered by the CNN-based technique and the systolic and diastolic blood pressure are well predicted with relative errors less than 6%. Overall, there was a good agreement between the CNN-based results which could be obtained in real-time, and the inversion-based quantities for all cases.

This confirms that the proposed method is really a good alternative to the inversion procedure, in particular in clinical applications where measurements are usually available at only a few locations and where faster diagnosis is always requested. In this sense, this approach paves the way for complete patient-specific vascular modeling, in which available uncertain clinical measurements (such as time series blood flow rates in the arteries) could be used to estimate (or tune) various uncertain parameters needed for the patient-specific simulations. The methodology has potential clinical applications such as providing reliable predictions of blood pressure in any artery of interest or in patient-specific treatment planning.

Finally, this thesis demonstrates, on the one hand, how computer simulations can provide an opportunity to obtain data currently unobtainable by other modalities and with essentially no risk for patients, and on the other hand, how the potential of ML, currently under-exploited could revolutionize biomedical research in the future.

7.2 Limitations and discussions

Several conclusions and limitations of this work have already been drawn and clarified in the course of this manuscript. It seems important to mention and discuss here two of them which are mainly of a modeling, methodological, and clinical aspect. Other limitations will be discussed along with the perspectives.

7.2.1 Use of a 0D model

As discussed in chapter 4, section 4.7, the model used in this work, namely the 0D model considers as uniform the distribution of fundamental variables (pressure, volume, and flow rate) at any instant in time and therefore, it does not account for the spatial variation of these variables and is only suitable for examination of global distributions of pressure, flow rate, and blood volume in the arterial network and for specific physiological conditions [91, 92]. Phenomena, such as wave transmission, wave travel, and wave reflections, cannot be or are not studied [326]. Moreover, as exposed in chapter 4, section 4.7 and chapter 6, section 6.7, several studies have pointed out the drawbacks of the lumped models and have proposed more complex and sophisticated models to better capture the characteristics of blood flow [91, 316, 327, 328, 329]. So we are aware that more sophisticated physical models can be considered. However, as the inversion procedure under consideration here employs a 0D model, to be able to properly examine the robustness and accuracy of the procedure and to make a fair comparison with ML-based technique results, it was necessary that uncertainty and sensitivity analysis as well as ML approach encapsulate the same physics as the one used in the inversion procedure. This is why the same cardiovascular model and arterial network as in the inversion procedure were considered in this work. Consequently, the use of a 0D model here should not be viewed as a deliberate choice but a constraint. Moreover, and this is the most important point, as already outlined in chapter 4, section 4.7, the region studied being compact, (the size of the cerebral network [of order 10 cm] is small compared with

typical pressure wavelength [of order a few meters]), assuming that all the quantities are homogeneous over each segment (viz, making the 0D assumption at the scale of each element), is acceptable. Of course, it would not be acceptable if this study was dealing with the arterial network of the whole body.

Furthermore, discussions on the choice of the appropriate model for any cardiovascular study are endless since as outlined by Shi et al. [91], each model is developed to achieve specific research purposes in each individual study. There is no universally optimal model that suits every application. Therefore, researchers must decide what level of model sophistication is most suitable in their specific studies and ensure that the complexity of their models fits the purposes of the studies. In the same vein, Kokalari et al. [92] claimed that a model cannot be named "bad" or "good", but its quality should be evaluated depending on the concrete purpose of its use. An over-simplified model, for example, will produce inadequate accuracy in the study. However, this does not mean that more complex model will always produce more accurate results [91]. In any case, as this work has attempted to accomplish more or less successfully, the development of a trustful CVS model should include the abstraction of the real system, experimental simulations, the comparison of the simulation results with the real system and the reduction of the eventual inconsistencies [92].

7.2.2 Blood pressure in the global sensitivity analysis

In the global sensitivity analysis approach performed in chapter 5 to identify the most important parameters in the inversion procedure, only blood flow rate was used as the patient-specific model output. Blood pressure was not considered at all. However, as outlined in chapter 4, section 4.5.3, about the blood flow model output with respect to a change in boundary conditions, simulated blood flow signals can match target signals while blood pressure is miscalculated. Therefore, sensitivity analysis carried out in chapter 5 using only blood flow as model output provides only the parameters

exerting large effect on the blood flow and not necessarily the most important parameters for blood pressure variations. However, it is worth pointing out that global sensitivity analysis conducted in chapter 5 was motivated by the availability of the patient-specific data. Indeed, only measured blood flow rates in the patient-specific internal carotid arteries were obtained from medical images. No time series pressure measurement was available. This lack of measured pressure data prevents from performing sensitivity analysis including the model pressure results since it was impossible to compare during the inversion procedure the estimated time series blood pressure to an non-existent patient-specific time series pressure. Unfortunately, this could not be avoided as measuring the pressure in these arteries was impossible. However, we have some confidence on the results because the flow rate and pressure are physically linked through the model. Consequently, in the absence of these pressure measurements and by virtue of the intrinsic link between blood flow and pressure, the global sensitivity analysis only based on the time series blood flow rates makes full sense. However, several improvements could be made in order to make the models and approaches adopted in this work more accurate.

7.3 Perspectives

Following this work, many possibilities for further research directions and future work can emerge. These mainly include modeling issues and clinical purposes, which could be addressed in the future.

7.3.1 Proper setting of model parameters

Since each lumped-parameter model of the CVS is generally a high level abstraction of the circulatory system and one model component often represents several complex anatomical structures, proper setting of model parameters is a key issue to be carefully addressed in the future [91]. For example, as highlighted in Section 4.4 of chapter 4, some model parameters used in this work were adopted from literature. This could lead

to inconsistent model parameters and call into question the consistency and reliability of the model. Further effort to properly set the model parameters would be therefore necessary in the future before the model used here can be considered as fully validated.

7.3.2 Multiscale modeling

As discussed above in chapter 2, Section 2.1 and in Section 7.2.1 of the current chapter, each of physical models of the CVS existing in the literature has its own scope [426, 427, 428]. The lumped-parameter model used in this work has its advantages, such as low computational cost and suitability to represent the global properties of the arterial network, but also some limitations, including poor accuracy, incapability to study the wave transmission phenomena, the effect of local vessel changes and the changes in the blood flow distributions. High-dimensional models, which are commonly used to describe in detail the local properties of the arterial network have also their strengths and weaknesses: they can provide high-precision simulation but with greater complexity. Therefore, to fully exploit the potential, the advantages of different dimensional models as in several previous works [305, 327, 328, 329, 330, 429, 430, 431], further effort is necessary in the future either to replace the simple 0D model used here with a more complex model (high-dimensional model) as in [108, 109, 432] or to couple models of various different scales of desired vascular structures [138, 433, 434, 435, 436, 437, 438] with the lumped parameter model used in this thesis. For instance, a collaboration with Dr. Emmanuelle LE BARS (*CHRU Gui de Chauliac, Montpellier, France*) could be envisaged in the near future to couple the model available at *CHRU Gui de Chauliac* including microcirculation with the model used in this work in order to validate first estimated cerebral blood pressure. Such multiscale models could clearly be powerful tools for providing potential applications in clinical practice.

7.3.3 Integration of other circulations and other body systems

The cerebral circulation mainly investigated in this thesis does not work in isolation. Other circulations, such as the systemic circulation or coronary circulation could be incorporated into the model used in this work in order to have a model of global CVS and to contribute to the apprehension of some pathologies, like myocardial infarction or stroke. Such a formulation of the closed loop cardiovascular model with the inclusion of a heart and venous system, will probably reduce the uncertainty coming from the in-output boundary conditions.

However, even if we manage to have a model of global CVS, there will still be a long way to go since the CVS does not work in isolation either. Indeed, there is a close interaction between the CVS with other systems such as the nervous system, the respiratory system, and the digestive system whose the effects are usually ignored in physical modeling of the CVS. However, these systems significantly influence the CVS [439, 440]. Study of their coupled reactions as in [441, 442, 443, 444, 445, 446], will bring the 0D cardiovascular model used in this work to a higher level of accuracy and will enormously improve the understanding of the patient-specific complex physiological and pathological mechanisms as in [447, 448]. Further effort should be continued in this direction in the future and this is a field with very promising prospects.

7.3.4 Need for extensive tests with the ML-based pressure estimator

Regarding the real-time blood pressure estimation in cerebral arteries using ML, the implicit CNN was tested successfully but only with one patient-specific clinical data (the same healthy subject as in the inversion procedure for the purpose of comparing the predicted blood pressure values in both cases). It is clear that this approach can play an increasingly key role in the development of medical devices but further tests could be carried out on a larger panel of subjects in order to confirm first the results presented in this thesis and before undertaking any actual clinical use. It could also be interesting to test the ML-based blood pressure estimator with data collected from

patients with vascular pathologies in order to verify whether the blood flow and blood pressure relationships established in each artery of interest in the healthy subject in this work are also observed in these patients. In addition, these patients may have other relationships between blood pressure and blood flow than those observed in the healthy subject. In any case, more extensive patient-specific tests and validation are necessary for the perspective of clinical applications before definite and final conclusions can be drawn with respect to the ML-based blood pressure estimator.

7.3.5 Other future directions

Other challenging problems to be solved in the future are listed as follows:

- As reported in Section 7.2.2, the lack of patient-specific time series blood pressure led us to use only time series blood flow as the model output for the global sensitivity analysis in chapter 5. However, the patient-specific measured SBP-DBP values being available, sensitivity analysis could be carried out using both these pressures information and the time series blood flow rates as the model output in the future.

- Future work will also involve investigation on modeling assumption used for the global sensitivity analysis in chapter 5. As a matter of fact, the linear relation assumed between the 21 estimated model parameters and the discrepancy between the simulated and the measured flow rate does not reflect the actual function of the hemodynamics in the blood flow model. Such a assumption could therefore be aberrant since hemodynamics is far from linear. However, this kind of assumptions is commonly used to approximate various highly complex phenomena not sufficiently known in order to get some insight into how the investigated phenomena actually work. Sensibility analysis including a realistic relationship between the estimated parameters and the difference between simulated and measured time series blood flow rate could be considered in the future.

- The blood flow model accuracy for patient-specific simulations is dependent on the quality of input data from imaging techniques. To fully exploit the potential of the ML-based blood pressure estimator, an improvement in patient-specific data acquisition procedures would be necessary in the future.

- As stated by Shi et al. [91], previous studies have suggested a number of cardiovascular effects that need to be specially addressed in the 0D cardiovascular modeling, including ventricular interaction, effect of pericardial, atrial-ventricular interaction, auto-regulation in some local circulation loops, auxiliary pumping action to blood flow caused by peripheral muscle contraction, venous valve in some vessels such as in the lower extremity etc. Although intuitively correct, these effects have received very limited validations to date, due to the major challenge in isolating each of these effects from the overall cardiovascular response (which is often a coupled interaction among the different organs) and also to the restrictions in *in vivo* measurements [91]. Therefore, further effort should be made in the future to quantify the relative importance of each of these effects in the response of the blood flow model used in this work.

It is thus clear that several improvements could be brought to make the model used and the approach proposed in this work more accurate. Nevertheless, this work already provides relevant results and useful information for noninvasively and quickly estimating patient-specific blood pressure in cerebral arteries, which could be valuable in the diagnosis and treatment to tackle cerebral aneurysms.

Bibliography

1. Ralph L Sacco, Scott E Kasner, Joseph P Broderick, Louis R Caplan, JJ Connors, Antonio Culebras, Mitchell SV Elkind, Mary G George, Allen D Hamdan, Randall T Higashida, et al. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the american heart association/american stroke association. *Stroke*, 44(7):2064–2089, 2013.

2. Eugene V Golanov, Evgeniy I Bovshik, Kelvin K Wong, Robia G Pautler, Chase H Foster, Richard G Federley, Jonathan Y Zhang, James Mancuso, Stephen TC Wong, and Gavin W Britz. Subarachnoid hemorrhage–induced block of cerebrospinal fluid flow: role of brain coagulation factor iii (tissue factor). *Journal of Cerebral Blood Flow & Metabolism*, 38(5):793–808, 2018.

3. Masayo Koide, Inna Sukhotinsky, Cenk Ayata, and George C Wellman. Subarachnoid hemorrhage, spreading depolarizations and impaired neurovascular coupling. *Stroke research and treatment*, 2013, 2013.

4. Joshua B Bederson, E Sander Connolly Jr, H Hunt Batjer, Ralph G Dacey, Jacques E Dion, Michael N Diringer, John E Duldner Jr, Robert E Harbaugh, Aman B Patel, and Robert H Rosenwasser. Guidelines for the management of aneurysmal subarachnoid hemorrhage: a statement for healthcare professionals from a special writing group of the stroke council, american heart association. *Stroke*, 40(3):994–1025, 2009.

5. Robert D Martin. *Subarachnoid Hemorrhage: Neurological Care and Protection*, volume 127. Springer, 2019.

6. Vincent Costalat, Mathieu Sanchez, Dominique Ambard, L Thines, Nicolas Lonjon, Franck Nicoud, H Brunel, JP Lejeune, Henri Dufour, P Bouillot, et al.

Biomechanical wall properties of human intracranial aneurysms resected following surgical clipping (irras project). *Journal of biomechanics*, 44(15):2685–2691, 2011.

7. Mathieu Sanchez, Dominique Ambard, Vincent Costalat, Simon Mendez, Franck Jourdan, and Franck Nicoud. Biomechanical assessment of the individual risk of rupture of cerebral aneurysms: a proof of concept. *Annals of biomedical engineering*, 41(1):28–40, 2013.

8. Christopher L Taylor, Zhong Yuan, Warren R Selman, Robert A Ratcheson, and Alfred A Rimm. Cerebral arterial aneurysm formation and rupture in 20,767 elderly patients: hypertension and other risk factors. *Journal of neurosurgery*, 83 (5):812–819, 1995.

9. Alexander Keedy. An overview of intracranial aneurysms. *McGill Journal of Medicine: MJM*, 9(2):141, 2006.

10. Yoshiteru Tada, Kosuke Wada, Kenji Shimada, Hiroshi Makino, Elena I Liang, Shoko Murakami, Mari Kudo, Keiko T Kitazato, Shinji Nagahiro, and Tomoki Hashimoto. Roles of hypertension in the rupture of intracranial aneurysms. *Stroke*, 45(2):579–586, 2014.

11. D Mohan, V Munteanu, T Coman, and AV Ciurea. Genetic factors involves in intracranial aneurysms–actualities. *Journal of medicine and life*, 8(3):336, 2015.

12. Rajnesh Lal, Franck Nicoud, E Bars, Jérémy Deverdun, F Molino, V Costalat, and B Mohammadi. Non invasive blood flow features estimation in cerebral arteries from uncertain medical data. *Annals of Biomedical Engineering*, 45:1–18, 08 2017. doi: 10.1007/s10439-017-1904-7.

13. J.G. Betts, P. Desaix, E.W. Johnson, J.E. Johnson, O. Korol, D. Kruse, B. Poe, OpenStax College, J. Wise, M.D. Womble, et al. *Anatomy & Physiology*. Open Textbook Library. OpenStax College, Rice University, 2013. ISBN 9781938168130. URL `https://books.google.fr/books?id=dvVgngEACAAJ`.

14. E.N. Marieb and K. Hoehn. *Human Anatomy & Physiology*. Pearson, 9 edition, 2013. ISBN 9780321743268. URL https://books.google.fr/books?id=MythtQAACAAJ.

15. A Noble, R Johnson, A Thomas, and P Bass. *The Cardiovascular System. Basic science and clinical Conditions*. Systems of the Body Series. Elsiever Churchill Livingstone, London, 2 edition, 02 2010. ISBN 9780702050824.

16. L.S. Lilly. *Braunwald's heart disease: a textbook of cardiovascular medicine*, volume 1. Elsevier Health Sciences, 2012.

17. D. Shier, J. Butler, and R. Lewis. *Hole's Human Anatomy*. McGraw-Hill, 1996.

18. Dorland. *Dorland's Illustrated Medical Dictionary E-Book*. Dorland's Medical Dictionary. Elsevier Health Sciences, 32 edition, 2011. ISBN 9781455709854. URL https://books.google.fr/books?id=mNACisYwbZoC.

19. TG Myers, Vicent Ribas Ripoll, Anna Sáez de Tejada Cuenca, Sarah L Mitchell, and Mark J McGuinness. Modelling the cardiovascular system for assessing the blood pressure curve. *Mathematics-in-industry case studies*, 8(1):2, 2017.

20. J Rodney Levick. *An introduction to cardiovascular physiology*. Butterworth-Heinemann, 2013.

21. Bernard Edward Matthews, M. Elizabeth Rogers, Michael Francis Oliver, John Higgins, Gloria Lotha, John M. Cunningham, Richard Pallardy, Kara Rogers, and Marco Sampaolo. Circulatory system. *Encyclopædia Britannica*, Aug 2019. URL <https://www.britannica.com/science/circulatory-system>. [Online; accessed 3-May-2020].

22. Institute for Quality and Efficiency in Health Care (IQWiG). How does the blood circulatory system work?, March 2010. URL https://www.informedhealth.org/how-does-the-blood-circulatory-system-work.2097.en.html. [Online; updated 2019 January 31].

23. L.R. Johnson and J.H. Byrne. *Essential Medical Physiology*. Lippincott-Raven, 1998. ISBN 9780397584017. URL https://books.google.fr/books?id=i6JpAAAAMAAJ.

24. K.S. Saladin. *Anatomy and Physiology: The Unity of Form and Function*. McGraw-Hill College, 1999. ISBN 9780076015436. URL https://books.google.ne/books?id=pm8MAAAACAAJ.

25. Barclay Tim. Cardiovascular system, April 2020. URL https://www.innerbody.com/image/cardov.html. [Online; accessed 3-May-2020].

26. Wikipedia contributors. Circulatory system — Wikipedia, the free encyclopedia, 2020. URL https://en.wikipedia.org/w/index.php?title=Circulatory_system&oldid=952074374. [Online; accessed 4-May-2020].

27. Christophe Chnafa. *Using image-based large-eddy simulations to investigate the intracardiac flow and its turbulent nature*. PhD thesis, University of Montpellier, 2014.

28. J. Mazumdar. *Biofluid Mechanics*. World Scientific, 1992. ISBN 9789810209278. URL https://books.google.td/books?id=bokW4-bVk9oC.

29. Gerard J Tortora, Bryan H Derrickson, Brendan Burkett, Danielle Dye, Julie Cooke, Tara Diversi, Mark McKean, Rebecca Mellifont, Latika Samalia, and Gregory Peoples. *Principles of anatomy and physiology*. John Wiley & Sons, United States, first asia - pacific edition edition, 2016. ISBN 0730314626. Includes bibliographical references and index.

30. Jagannath Mazumdar. *Biofluid Mechanics*. WORLD SCIENTIFIC, Singapore, 2 edition, 2016. doi: 10.1142/9754. URL https://www.worldscientific.com/doi/abs/10.1142/9754.

31. Arthur C Guyton, John Edward Hall, et al. *Textbook of medical physiology*, volume 548. Saunders Philadelphia, 1986.

32. John E Hall. *Guyton and Hall textbook of medical physiology e-Book*. Elsevier Health Sciences, 2010.

33. John E Hall. *Guyton and Hall Textbook of Medical Physiology, Jordanian Edition E-Book*. Elsevier, 2016.

34. United States National Institutes of Health, National Cancer Institute. Seer training modules. classification & structure of blood vessels, 2020. URL <https://training.seer.cancer.gov/anatomy/cardiovascular/blood/classification.html>. [Online; accessed 3-May-2020].

35. Boundless. Boundless Anatomy and Physiology. Cardiovascular System: Blood Vessels - Blood Vessel Structure and Function, 06 2019. URL <https://courses.lumenlearning.com/boundless-ap/chapter/blood-vessel-structure-and-function/>. [online; accessed 24-April-2020].

36. Mark L. Entman, Stanley W. Jacob, Michael Francis Oliver, John Higgins, Gloria Lotha, John M. Cunningham, Richard Pallardy, Kara Rogers, and Marco Sampaolo. Human cardiovascular system. *Encyclopædia Britannica*, Feb 2020. URL <https://www.britannica.com/science/human-cardiovascular-system>. [Online; accessed 3-May-2020].

37. Kevin T Patton. *Anatomy & Physiology Laboratory Manual and E-Labs E-Book*. Elsevier Health Sciences, 2018.

38. Blausen Medical Communications, Inc. Blausen.com staff (2014), medical gallery of blausen medical 2014. *WikiJournal of Medicine*, 1(2), 2014. ISSN 2002-4436. doi: 10.15347/wjm/2014.010. URL https://app.dimensions.ai/details/publication/pub.1067741288andhttps://upload.

wikimedia.org/wikiversity/en/7/72/Blausen_gallery_2014.pdf. [Online; accessed 3-May-2020].

39. Boundless. Boundless Anatomy and Physiology. Cardiovascular System: Blood Vessels - Arteries, 06 2019. URL <https://courses.lumenlearning.com/boundless-ap/chapter/arteries/>. [online; accessed 24-April-2020].

40. Ali Ostadfar. Chapter 2 - macrocirculation system. In Ali Ostadfar, editor, *Biofluid Mechanics*, pages 61 – 86. Academic Press, 2016. ISBN 978-0-12-802408-9. doi: https://doi.org/10.1016/B978-0-12-802408-9.00002-8. URL http://www.sciencedirect.com/science/article/pii/B9780128024089000028.

41. R.E. Shadwick. Mechanical design in arteries. *Journal of Experimental Biology*, 202(23):3305–3313, 1999. ISSN 0022-0949. URL https://jeb.biologists.org/content/202/23/3305.

42. William D Tucker, Yingyot Arora, and Kunal Mahajan. Anatomy, blood vessels. In *StatPearls [Internet]*. Treasure Island (FL):StatPearls Publishing, 2019. URL https://www.ncbi.nlm.nih.gov/books/NBK470401/?log$=activity.

43. Wikipedia contributors. Muscular artery — Wikipedia, the free encyclopedia, 2020. URL https://en.wikipedia.org/w/index.php?title=Muscular_artery&oldid=945557490. [Online; accessed 24-April-2020].

44. A. Maton. *Human Biology and Health*. Prentice Hall science. Englewood Cliffs, New Jersey: Pearson Prentice Hall, 1993. ISBN 9780139811760. URL https://archive.org/details/humanbiologyheal00scho/page/6/mode/2up.

45. Wikipedia contributors. Arteriole — Wikipedia, the free encyclopedia, 2019. URL https://en.wikipedia.org/w/index.php?title=Arteriole&oldid=932012691. [Online; accessed 24-April-2020].

46. Paul Johns. Chapter 10 - stroke. In Paul Johns, editor, *Clinical Neuroscience*, pages 115 – 128. Churchill Livingstone, 2014. ISBN 978-0-443-10321-6. doi: https://doi.org/10.1016/B978-0-443-10321-6.00010-2. URL http://www.sciencedirect.com/science/article/pii/B9780443103216000102.

47. D.E. Haines. Chapter 8 - a survey of the cerebrovascular system. In Duane E. Haines and Gregory A. Mihailoff, editors, *Fundamental Neuroscience for Basic and Clinical Applications (Fifth Edition)*, pages 122 – 137.e1. Elsevier, fifth edition edition, 2018. ISBN 978-0-323-39632-5. doi: https://doi.org/10.1016/B978-0-323-39632-5.00008-6. URL http://www.sciencedirect.com/science/article/pii/B9780323396325000086.

48. Wikipedia contributors. Circle of willis — Wikipedia, the free encyclopedia, 2020. URL https://en.wikipedia.org/w/index.php?title=Circle_of_Willis&oldid=950739786. [Online; accessed 24-April-2020].

49. T. David and R.G. Brown. Chapter 7 - models of cerebrovascular perfusion. In Sid M. Becker and Andrey V. Kuznetsov, editors, *Transport in Biological Media*, pages 253 – 273. Elsevier, Boston, 2013. ISBN 978-0-12-415824-5. doi: https://doi.org/10.1016/B978-0-12-415824-5.00007-2. URL http://www.sciencedirect.com/science/article/pii/B9780124158245000072.

50. Zvonimir Vrselja, Hrvoje Brkic, Stefan Mrdenovic, Radivoje Radic, and Goran Curic. Function of circle of willis. *Journal of Cerebral Blood Flow & Metabolism*, 34(4):578–584, 2014. doi: 10.1038/jcbfm.2014.7. URL https://doi.org/10.1038/jcbfm.2014.7. PMID: 24473483.

51. Cagatay Uston. NEUROwords Dr. Thomas Willis' Famous Eponym: The Circle of Willis. *Journal of the History of the Neurosciences*, 14(1):16–21, 2005. doi: 10.1080/096470490512553. URL https://doi.org/10.1080/096470490512553. PMID: 15804755.

52. Stanley D. Johnsen. Book review: Clinical neuropathology. text and color atlas. by catherine haberland. demos publishing, new york, ny, 2007. *Journal of Child Neurology*, 23(4):466–466, 2008. doi: 10.1177/0883073808315421. URL https://doi.org/10.1177/0883073808315421.

53. David L. Felten, M. Kerry O'Banion, and Mary Summo Maida. 7 - vasculature. In David L. Felten, M. Kerry O'Banion, and Mary Summo Maida, editors, *Netter's Atlas of Neuroscience (Third Edition)*, pages 93 – 124. Elsevier, Philadelphia, third edition edition, 2016. ISBN 978-0-323-26511-9. doi: https://doi.org/10.1016/B978-0-323-26511-9.00007-2. URL http://www.sciencedirect.com/science/article/pii/B9780323265119000072.

54. A. Waugh and A. Grant. *Ross & Wilson Anatomy and Physiology in Health and Illness E-Book*. Elsevier Health Sciences, 2018. ISBN 9780702072840. URL https://books.google.fr/books?id=9w5kDwAAQBAJ.

55. Wanda G. Webb. 3 - organization of the nervous system ii. In Wanda G. Webb, editor, *Neurology for the Speech-Language Pathologist (Sixth Edition)*, pages 44 – 73. Mosby, sixth edition edition, 2017. ISBN 978-0-323-10027-4. doi: https://doi.org/10.1016/B978-0-323-10027-4.00003-8. URL http://www.sciencedirect.com/science/article/pii/B9780323100274000038.

56. Dale Purves, George J Augustine, David Fitzpatrick, William C Hall, Anthony-Samuel LaMantia, James O McNamara, and S Mark Williams. *Neuroscience*. Sunderland, MA: Sinauer Associates,Inc, fourth edition, 2007.

57. Thomas Willis. *Anatomy of the Brain and Nerves: Volumes 1 & 2*. McGill-Queen's Press-MQUP, 1965.

58. Thomas Willis, William Feindel, and Samuel Pordage. *The anatomy of the brain and nerves*. Montreal [Quebec] : McGill University Press, tercentenary edition, 1664-1964 edition, 2014. ISBN 9780773594401.

URL http://gateway.library.qut.edu.au/login?url=http://www.jstor.org/stable/10.2307/j.ctt1t88wg2. Feindel, William (editor) and Pordage, Samuel 1633-1691 (translator). Translation of Cerebri anatome, cui accessit Nervorum descriptio et usus.

59. William B. Lo and Harold Ellis. The Circle Before Willis: A Historical Account of the Intracranial Anastomosis. *Neurosurgery*, 66(1):7–18, 01 2010. ISSN 0148-396X. doi: 10.1227/01.NEU.0000362002.63241.A5. URL https://doi.org/10.1227/01.NEU.0000362002.63241.A5.

60. Charles Symonds. The circle of willis. *British medical journal*, 1(4906):119, 1955.

61. AWJ Hoksbergen, B Fulesdi, DA Legemate, and L Csiba. Collateral configuration of the circle of willis: transcranial color-coded duplex ultrasonography and comparison with postmortem anatomy. *Stroke*, 31(6):1346–1351, 2000.

62. RR Edelman, HP Mattle, GV O'Reilly, KU Wentz, Cheng Liu, and Bin Zhao. Magnetic resonance imaging of flow dynamics in the circle of willis. *Stroke*, 21(1): 56–65, 1990.

63. H Nornes. The role of the circle of willis in graded occlusion of the internal carotid artery in man. *Acta neurochirurgica*, 28(3):165–177, 1973.

64. K.V. Kardong. *Vertebrates: Comparative Anatomy, Function, Evolution*. McGraw-Hill, 4 edition, 2006. ISBN 9780071244572. URL https://books.google.fr/books?id=JNs-AQAAIAAJ. McGraw-Hill, ed. Original provenant de l'Université de Californie. Numérisé le 8-12-2010.

65. Michiel J de Boorder, Jeroen van der Grond, Alice J van Dongen, Catharina J M Klijn, L Jaap Kappelle, Peter P Van Rijk, and Jeroen Hendrikse. Spect measurements of regional cerebral perfusion and carbondioxide reactivity: correlation with cerebral collaterals in internal carotid artery occlusive disease. *Journal of neurology*, 253(10):1285–1291, October 2006. ISSN 0340-5354. doi: 10.

1007/s00415-006-0192-1. URL https://doi.org/10.1007/s00415-006-0192-1.

66. Kelly Skinner. 227 Best Interventional Radiology images, 04 2020. URL <https://www.pinterest.fr/kellysakacs/interventional-radiology/>. [online; accessed 24-April-2020].

67. Wikimedia Commons. File:circle of willis en.svg — wikimedia commons, the free media repository, 2019. URL https://training.seer.cancer.gov/citation.html. [Online; accessed 28-April-2020].

68. James P Keener and James Sneyd. *Mathematical physiology*, volume 1. Springer, 1998.

69. Maxime Cannesson, Cyril Besnard, Pierre G Durand, Julien Bohé, and Didier Jacques. Relation between respiratory variations in pulse oximetry plethysmographic waveform amplitude and arterial pulse pressure in ventilated patients. *Critical Care*, 9(5):R562, 2005.

70. Andrew Reisner, Phillip A Shaltis, Devin McCombie, and H Harry Asada. Utility of the photoplethysmogram in circulatory monitoring. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 108(5):950–958, 2008.

71. K Shelley and S Shelley. Clinical monitoring: practical applications for anesthesia and critical care, 2001.

72. Kirk H Shelley. Photoplethysmography: beyond the calculation of arterial oxygen saturation and heart rate. *Anesthesia & Analgesia*, 105(6):S31–S36, 2007.

73. William Joyce, Daniel W. White, Peter B. Raven, and Tobias Wang. Weighing the evidence for using vascular conductance, not resistance, in comparative cardiovascular physiology. *Journal of Experimental Biology*, 222(6), 2019. ISSN 0022-0949. doi: 10.1242/jeb.197426. URL https://jeb.biologists.org/content/222/6/jeb197426.

74. Timothy W. Secomb. Hemodynamics. *Comprehensive Physiology*, 6(2):975–1003, 04 2016. doi: 10.1002/cphy.c150038. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cphy.c150038.

75. Euroform Healthcare. Medical Physiology. Vascular Distensibility., 04 2019. URL <https://www.euroformhealthcare.biz/medical-physiology/vascular-distensibility.html>. [online; accessed 24-April-2020]; Last Updated on Wed, 04 Sep 2019.

76. Luc M. Van Bortel, Mirian J. Kool, and Harry A. Struijker Boudier. Effects of antihypertensive agents on local arterial distensibility and compliance. *Hypertension*, 26(3):531–534, 1995. doi: 10.1161/01.HYP.26.3.531. URL https://www.ahajournals.org/doi/abs/10.1161/01.HYP.26.3.531.

77. Michel E Safar. Pulse pressure in essential hypertension: clinical and therapeutical implications. *Journal of hypertension*, 7(10):769–776, 1989.

78. World Health Organization. Cardiovascular diseases (CVDs). The United Nations, 05 2017. URL <https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)>. [online; accessed 24-April-2020].

79. National Health Service. The United Kingdom. Cardiovascular disease, 2018. URL https://www.nhs.uk/conditions/cardiovascular-disease/. [Online; accessed 24-April-2020].

80. Rachel Hajar. Risk Factors for Coronary Artery Disease: Historical Perspectives. *Heart Views*, 18(3):109 – 114, 2017. ISSN 0976-5123. doi: 10.4103/HEARTVIEWS.HEARTVIEWS_106_17. URL https://pubmed.ncbi.nlm.nih.gov/29184622.

81. World Heart Fundation. Risk factors, 2020. URL <hhttps://www.world-heart-federation.org/resources/risk-factors/>. [online; accessed 24-April-2020].

82. Lusis Aldons J. Atherosclerosis. *Nature*, 407(6801):233–241, 09 2000. ISSN 0028-0836. doi: 10.1038/35025203. URL https://pubmed.ncbi.nlm.nih.gov/11001066. [Online; accessed 3-May-2020].

83. University of California San Francisco. Cardiac Surgery. Department of Surgery. Conditions & Procedures. Atherosclerosis, 04 2020. URL <https://cardiacsurgery.ucsf.edu/conditions--procedures/atherosclerosis.aspx>. [online; accessed 24-April-2020].

84. University of California San Francisco. Cardiac Surgery. Department of Surgery. Conditions & Procedures. Coronary Artery Disease, 04 2020. URL <https://cardiacsurgery.ucsf.edu/conditions--procedures/coronary-artery-disease.aspx>. [online; accessed 24-April-2020].

85. National Health Service. The United Kingdom. Atherosclerosis (arteriosclerosis), 2019. URL https://www.nhs.uk/conditions/atherosclerosis/. [Online; accessed 24-April-2020; last reviewed 2-May-2019].

86. Raheem Gul. *Mathematical modeling and sensitivity analysis of lumped-parameter model of the human cardiovascular system.* PhD thesis, Freien Universität Berlin, 2016. URL https://www.researchgate.net/publication/301852867_Mathematical_Modeling_and_Sensitivity_Analysis_of_Lumped-Parameter_Model_of_the_Human_Cardiovascular_System.

87. National Institutes of Health. National Heart, Lung, and Blood Institute (NHLBI). Atherosclerosis. *The United Nations, accessed 24 april 2020*, 2020. URL https://www.nhlbi.nih.gov/health-topics/atherosclerosis.

88. Columbia University. Department of Neurology. Cerebral Aneurysm, 2001. [Online; accessed 3-May-2020].

89. Brain Aneurysm Foundation. Brain aneurysm basics, 2020. URL https://bafound.org/about-brain-aneurysms/brain-aneurysm-basics/.

90. Wikimedia Commons. File:cerebral aneurysm nih.jpg — wikimedia commons, the free media repository, 2019. URL https://commons.wikimedia.org/w/index.php?title=File:Cerebral_aneurysm_NIH.jpg&oldid=383959349. [Online; accessed 2-May-2020].

91. Yubing Shi, Patricia Lawford, and Rodney Hose. Review of zero-d and 1-d models of blood flow in the cardiovascular system. *BioMedical Engineering OnLine*, 10 (1):33, Apr 2011. ISSN 1475-925X. doi: 10.1186/1475-925X-10-33. URL https://doi.org/10.1186/1475-925X-10-33.

92. Isidor Kokalari, Theodhor Karaja, and Maria Guerrisi. Review on lumped parameter method for modeling the blood flow in systemic arteries. *Journal of Biomedical Science and Engineering*, 6(1):92–99, 2013. doi: 10.4236/jbise.2013.61012. URL http://file.scirp.org/Html/12-9101565_27458.htm.

93. Shuran Zhou, Lisheng Xu, Liling Hao, Hanguang Xiao, Yang Yao, Lin Qi, and Yudong Yao. A review on low-dimensional physics-based models of systemic arteries: application to estimation of central aortic pressure. *Biomedical engineering online*, 18(1):41, 2019.

94. ALFIO Quarteroni, Andrea Manzoni, and Christian Vergara. The cardiovascular system: mathematical modelling, numerical algorithms and clinical applications. *Acta Numerica*, 26:365–590, 2017. doi: 10.1017/S0962492917000046.

95. Claudio De Lazzari and Marina Pirtskhalava. *Cardiovascular and Pulmonary Artificial Organs: Educational Training Simulators*. CNR Edizioni, 05

2017. ISBN 9788880802402. URL https://books.google.fr/books?id=l8DgswEACAAJ.

96. Darshan Doshi and Daniel Burkhoff. Cardiovascular simulation of heart failure pathophysiology and therapeutics. *Journal of cardiac failure*, 22(4):303–311, 2016.

97. Gianfranco Ferrari, Arianna Di Molfetta, Krzysztof Zieliński, and Libera Fresiello. Circulatory modelling as a clinical decision support and an educational tool. *Biomed Data J*, 1(3):45–50, 2015.

98. Massimo Capoccia, Silvia Marconi, Sanjeet Avtaar Singh, Domenico M Pisanelli, and Claudio De Lazzari. Simulation as a preoperative planning approach in advanced heart failure patients. a retrospective clinical analysis. *Biomedical engineering online*, 17(1):52, 2018.

99. S. Mendez, E. Gibaud, and F. Nicoud. An unstructured solver for simulations of deformable particles in flows at arbitrary reynolds numbers. *JCP*, 256:465–483, 2014.

100. Franck Nicoud, Christophe Chnafa, J Siguenza, Vladeta Zmijanovic, and Simon Mendez. Large-eddy simulation of turbulence in cardiovascular flows. In *Biomedical Technology*, pages 147–167. Springer, 2018.

101. Rodrigo Méndez Rojano, Simon Mendez, and Franck Nicoud. Introducing the procoagulant contact system in the numerical assessment of device-related thrombosis. *Biomechanics and modeling in mechanobiology*, 17(3):815–826, 2018.

102. S Mendez, C Chnafa, E Gibaud, J Sigüenza, V Moureau, and F Nicoud. Yales2bio: a computational fluid dynamics software dedicated to the prediction of blood flows in biomedical devices. In *5th International Conference on Biomedical Engineering in Vietnam*, pages 7–10. Springer, 2015.

103. Christophe Chnafa, Simon Mendez, and Franck Nicoud. Image-based simulations show important flow fluctuations in a normal left ventricle: What could be the implications? *Annals of biomedical engineering*, 44(11):3346–3358, 2016.

104. Pierre Taraconat, Jean-Philippe Gineys, Damien Isebe, Franck Nicoud, and Simon Mendez. Numerical simulation of deformable particles in a coulter counter. *International Journal for Numerical Methods in Biomedical Engineering*, 35(11): e3243, 2019. doi: 10.1002/cnm.3243. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3243`. e3243 cnm.3243.

105. Vladeta Zmijanovic, Simon Mendez, Vincent Moureau, and Franck Nicoud. About the numerical robustness of biomedical benchmark cases: Interlaboratory fda's idealized medical device. *International Journal for Numerical Methods in Biomedical Engineering*, 33(1):e02789, 2017. doi: 10.1002/cnm.2789. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2789`. e02789 cnm.2789.

106. Thomas Puiseux, Anou Sewonu, Olivier Meyrignac, Hervé Rousseau, Franck Nicoud, Simon Mendez, and Ramiro Moreno. Reconciling pc-mri and cfd: An in-vitro study. *NMR in Biomedicine*, 32(5):e4063, 2019. doi: 10.1002/nbm.4063. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/nbm.4063`. e4063 nbm.4063.

107. J. Sigüenza, S. Mendez, D. Ambard, F. Dubois, F. Jourdan, R. Mozul, and F. Nicoud. Validation of an immersed thick boundary method for simulating fluid-structure interactions of deformable membranes. *Journal of Computational Physics*, 322:723–746, October 2016. doi: 10.1016/j.jcp.2016.06.041.

108. S Boujena, Oualid Kafi, and N El Khatib. A 2D mathematical model of blood flow and its interactions in an atherosclerotic artery. *Mathematical Modelling of Natural Phenomena*, 9(6):46–68, 2014.

109. Arthur R Ghigo, Jose-Maria Fullana, and Pierre-Yves Lagrée. A 2D nonlinear multiring model for blood flow in large elastic arteries. *Journal of Computational Physics*, 350:136–165, 2017.

110. Xinzhou Xie, Minwen Zheng, Didi Wen, Yabing Li, and Songyun Xie. A new CFD based non-invasive method for functional diagnosis of coronary stenosis. *Biomedical engineering online*, 17(1):36, 2018.

111. Alejandro Lopez-Perez, Rafael Sebastian, and Jose M Ferrero. Three-dimensional cardiac computational modelling: methods, features and applications. *Biomedical engineering online*, 14(1):35, 2015.

112. Nicolaas Westerhof, Frederik Bosman, Cornelis J. De Vries, and Abraham Noordergraaf. Analog studies of the human systemic arterial tree. *Journal of Biomechanics*, 2(2):121–143, 1969. ISSN 0021-9290. doi: https://doi.org/10.1016/0021-9290(69)90024-4. URL `http://www.sciencedirect.com/science/article/pii/0021929069900244`.

113. Thomas JR Hughes and J Lubliner. On the one-dimensional theory of blood flow in the larger vessels. *Mathematical Biosciences*, 18(1):161–170, 1973.

114. John Wallace Lambert. *Fluid flow in a nonrigid tube*. PhD thesis, Purdue University, 1956. Doctoral Dissertation.

115. Mette S Olufsen, Charles S Peskin, Won Yong Kim, Erik M Pedersen, Ali Nadim, and Jesper Larsen. Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions. *Annals of biomedical engineering*, 28(11):1281–1299, 2000.

116. Annemette Sofie Olufsen. *Modeling the arterial system with reference to an anesthesia simulator*. PhD thesis, Roskilde University, 1998.

117. Majid Abdollahzade, Chang-Sei Kim, Nima Fazeli, Barry A Finegan, M Sean McMurtry, and Jin-Oh Hahn. Data-driven lossy tube-load modeling of arterial tree: In-human study. *Journal of biomechanical engineering*, 136(10), 2014.

118. Mingwu Gao, Hao-Min Cheng, Shih-Hsien Sung, Chen-Huan Chen, Nicholas Bari Olivier, and Ramakrishna Mukkamala. Estimation of pulse transit time as a function of blood pressure using a nonlinear arterial tube-load model. *IEEE Transactions on Biomedical Engineering*, 64(7):1524–1534, 2016.

119. Roberto Burattini and Silvia Natalucci. Complex and frequency-dependent compliance of viscoelastic windkessel resolves contradictions in elastic windkessels. *Medical Engineering & Physics*, 20(7):502 – 514, 1998. ISSN 1350-4533. doi: https://doi.org/10.1016/S1350-4533(98)00055-1. URL `http://www.sciencedirect.com/science/article/pii/S1350453398000551`.

120. A. P. Avolio. Multi-branched model of the human arterial system. *Medical and Biological Engineering and Computing*, 18(6):709–718, Nov 1980. ISSN 1741-0444. doi: 10.1007/BF02441895. URL `https://doi.org/10.1007/BF02441895`.

121. Vuk Milišić and Alfio Quarteroni. Analysis of lumped parameter models for blood flow simulations and their relation with 1D models. *ESAIM: Mathematical Modelling and Numerical Analysis*, 38(4):613–632, 2004.

122. Abraham Noordergraaf, Pieter D. Verdouw, and Herman B.K. Boom. The use of an analog computer in a circulation model. *Progress in Cardiovascular Diseases*, 5(5):419 – 439, 1963. ISSN 0033-0620. doi: https://doi.org/10.1016/S0033-0620(63)80009-2. URL `http://www.sciencedirect.com/science/article/pii/S0033062063800092`.

123. Canek Moises Luna Phillips. *Simple lumped parameter model of the cardiovascular system, A*. PhD thesis, Colorado State University. Libraries, 2011.

124. SHIM Eun Bo, SAH Jong Youb, and YOUN Chan Hyun. Mathematical modeling of cardiovascular system dynamics using a lumped parameter method. *The Japanese Journal of Physiology*, 54(6):545–553, mar 2005. ISSN 0021521X. doi: 10.2170/jjphysiol.54.545. URL https://ci.nii.ac.jp/naid/10014467893/en/.

125. Yubing Shi. *Lumped-parameter modelling of cardiovascular system dynamics under different healthy and diseased conditions*. PhD thesis, University of Sheffield, 2013.

126. Rajnesh Lal. *Data assimilation and uncertainty quantification in cardiovascular biomechanices*. PhD thesis, University of Montpellier, 2017.

127. S Malatos, A Raptis, and M Xenos. Advances in low-dimensional mathematical modeling of the human cardiovascular system. *J Hypertens Manag*, 2(2):1–10, 2016.

128. Otto Frank. Die grundform des arteriellen pulses. *Z Biol*, 37(483-526):459, 1899.

129. S Hales. 1733 statical essays: Containing haemostaticks, or an account of some hydraulick and hydrostatical experiments made on the blood and blood vessels of animals, 1733.

130. S Hales. Statical essays: Containing haemostaticks, vol. ii. *London, England: Innys and Manby*, 1733, 1997.

131. Christopher M. Quick, David S. Berger, and Abraham Noordergraaf. Apparent arterial compliance. *American Journal of Physiology-Heart and Circulatory Physiology*, 274(4):H1393–H1403, 1998. doi: 10.1152/ajpheart.1998.274.4.H1393. URL https://doi.org/10.1152/ajpheart.1998.274.4.H1393. PMID: 29586011.

132. Nicolaas Westerhof, GIJS Elzinga, and P Sipkema. An artificial arterial system for pumping hearts. *Journal of applied physiology*, 31(5):776–781, 1971.

133. G Landes. Einige untersuchungen an elektrischen analogieschaltungen zum kreitslaufsystem. *Z Biol*, 101:418–429, 1943.

134. D Burkhoff, J Alexander Jr, and J Schipke. Assessment of windkessel as a model of aortic input impedance. *American Journal of Physiology-Heart and Circulatory Physiology*, 255(4):H742–H753, 1988.

135. KH Wesseling, JR Jansen, JJ Settels, and JJ Schreuder. Computation of aortic flow from pressure in humans using a nonlinear, three-element model. *Journal of applied physiology*, 74(5):2566–2573, 1993.

136. WARREN K Laskey, H GRAHAM Parker, VICTOR A Ferrari, WILLIAM G Kussmaul, and ABRAHAM Noordergraaf. Estimation of total systemic arterial compliance in humans. *Journal of Applied Physiology*, 69(1):112–119, 1990.

137. N Stergiopulos, DF Young, and TR Rogge. Computer simulation of arterial flow with applications to arterial and aortic stenoses. *Journal of biomechanics*, 25(12): 1477–1488, 1992.

138. Luca Formaggia and Alessandro Veneziani. Reduced and multiscale models for the human cardiovascular system. *Lecture notes VKI lecture series*, 7, 2003.

139. R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, 03 1960. ISSN 0021-9223. doi: 10.1115/1.3662552. URL `https://doi.org/10.1115/1.3662552`.

140. Wikipedia contributors. Kalman filter — Wikipedia, the free encyclopedia, 2020. URL `https://en.wikipedia.org/w/index.php?title=Kalman_filter&oldid=960685929`. [Online; accessed 20-April-2020].

141. Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans (1978–2012)*, 99(C5):10143–10162, 1994.

142. Ricardo Todling and Stephen E Cohn. Suboptimal schemes for atmospheric data assimilation based on the kalman filter. *Monthly Weather Review*, 122(11):2530–2557, 1994.

143. Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45, 1960.

144. Andrew H Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

145. Arthur Gelb. *Applied optimal estimation*. MIT press, 1974.

146. Simon J Julier and Jeffrey K Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.

147. Gerrit Burgers, Peter Jan van Leeuwen, and Geir Evensen. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, 126(6):1719–1724, 06 1998. ISSN 0027-0644. doi: 10.1175/1520-0493(1998)126⟨1719:ASITEK⟩2.0.CO;2. URL https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

148. Stephen J Thomas, JP Hacker, and JL Anderson. A robust formulation of the ensemble kalman filter. *Quarterly Journal of the Royal Meteorological Society*, 135 (639):507–521, 2009.

149. Matthias Katzfuss, Jonathan R. Stroud, and Christopher K. Wikle. Understanding the ensemble kalman filter. *The American Statistician*, 70(4):350–357, 2016. doi: 10.1080/00031305.2016.1141709. URL https://doi.org/10.1080/00031305.2016.1141709.

150. John W Tukey. The future of data analysis. *The annals of mathematical statistics*, 33(1):1–67, 1962.

151. Jan Mandel. A brief tutorial on the ensemble kalman filter. *arXiv preprint arXiv:0901.3725*, 2009.

152. Susan Dunne and Dara Entekhabi. An ensemble-based reanalysis approach to land data assimilation. *Water Resources Research*, 41(2), 2005.

153. Rajnesh Lal, Bijan Mohammadi, and Franck Nicoud. Data assimilation for identification of cardiovascular network characteristics. *International Journal for Numerical Methods in Biomedical Engineering*, 33(5):e2824, may 2017. doi: 10.1002/cnm.2824. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2824`. e2824 cnm.2824.

154. Geir Evensen. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.

155. Jeffrey S Whitaker and Thomas M Hamill. Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130(7):1913–1924, 2002.

156. Zheqi Shen and Youmin Tang. A modified ensemble kalman particle filter for non-gaussian systems with nonlinear measurement functions. *Journal of Advances in Modeling Earth Systems*, 2015.

157. Peter L Houtekamer and Herschel L Mitchell. A sequential ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1):123–137, 2001.

158. Jaison Thomas Ambadan and Youmin Tang. Sigma-Point Kalman Filter Data Assimilation Methods for Strongly Nonlinear Systems. *Journal of the Atmospheric Sciences*, 66(2):261–285, 02 2009. ISSN 0022-4928. doi: 10.1175/2008JAS2681.1. URL `https://doi.org/10.1175/2008JAS2681.1`.

159. Youmin Tang, Jaison Ambandan, and Dake Chen. Nonlinear measurement function in the ensemble kalman filter. *Advances in Atmospheric Sciences*, 31 (3):551–558, 2014.

160. Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global Sensitivity Analysis. The Primer*. John Wiley & Sons, Ltd, 2008. ISBN 9780470725184. doi: 10.1002/9780470725184. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470725184.

161. Hermann G. Matthies. Quantifying uncertainty: Modern computational representation of probability and applications. In Adnan Ibrahimbegovic and Ivica Kozar, editors, *Extreme Man-Made and Natural Hazards in Dynamics of Structures*, pages 105–135, Dordrecht, 2007. Springer Netherlands. ISBN 978-1-4020-5656-7.

162. Armen Der Kiureghian and Ove Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105 – 112, 2009. ISSN 0167-4730. doi: https://doi.org/10.1016/j.strusafe.2008.06.020. URL http://www.sciencedirect.com/science/article/pii/S0167473008000556. Risk Acceptance and Risk Communication.

163. Wikipedia contributors. Uncertainty quantification — Wikipedia, the free encyclopedia, 2020. URL https://en.wikipedia.org/w/index.php?title=Uncertainty_quantification&oldid=955461996. [Online; accessed 1-May-2020].

164. Marc C. Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001. doi: 10.1111/1467-9868.00294. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00294.

165. John Milburn McFarland. *Uncertainty analysis for computer simulations through validation and calibration*. PhD thesis, Vanderbilt University, Nashville, TN 37235, 2008.

166. Andrea Saltelli. Sensitivity analysis for importance assessment. *Risk Analysis*, 22(3):579–590, 2002. doi: 10.1111/0272-4332.00040. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/0272-4332.00040.

167. Andrea Saltelli, Stefano Tarantola, Francesca Campolongo, and Marco Ratto. *Sensitivity analysis in practice: a guide to assessing scientific models*, volume 1. Wiley Online Library, 2004.

168. Wikipedia contributors. Sensitivity analysis — Wikipedia, the free encyclopedia, 2020. URL https://en.wikipedia.org/w/index.php?title=Sensitivity_analysis&oldid=962434380. [Online; accessed 21-April-2020].

169. Robert Bailis, Majid Ezzati, and Daniel M. Kammen. Mortality and greenhouse gas impacts of biomass and petroleum energy futures in africa. *Science*, 308(5718): 98–103, 2005. ISSN 0036-8075. doi: 10.1126/science.1106881. URL https://science.sciencemag.org/content/308/5718/98.

170. David M Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32(2):135–154, 1994. doi: 10.1007/BF00547132.

171. Zhike Zi. Sensitivity analysis approaches applied to systems biology models. *IET systems biology*, 5(6):336–346, 2011.

172. J Elliott Campbell, Gregory R Carmichael, T Chai, M Mena-Carrasco, Y Tang, DR Blake, NJ Blake, Stephanie A Vay, G James Collatz, I Baker, et al. Photosynthetic control of atmospheric carbonyl sulfide during the growing season. *Science*, 322(5904):1085–1088, 2008.

173. James M Murphy, David MH Sexton, David N Barnett, Gareth S Jones, Mark J Webb, Matthew Collins, and David A Stainforth. Quantification of modelling

uncertainties in a large ensemble of climate change simulations. *Nature*, 430(7001): 768–772, 08 2004. URL https://doi.org/10.1038/nature02771.

174. Veronica Czitrom. One-factor-at-a-time versus designed experiments. *The American Statistician*, 53(2):126–131, 1999. ISSN 00031305. URL http://www.jstor.org/stable/2685731.

175. Andrea Saltelli, K Chan, and E.Marian Scott. *Sensitivity Analysis*, volume 134 of *Probability and statistics series*. John and Wiley & Sons, New York, 01 2000. ISBN 0471998923.

176. Andrea Saltelli, Stefano Tarantola, and KP-S Chan. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56, 1999.

177. Max D Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.

178. I. M. Sobol. On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe modelirovanie*, 2(2):112–118, 1990. URL http://mi.mathnet.ru/mm2320.

179. I.M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1:407–414, 1993.

180. Devanshi Dhall, Ravinder Kaur, and Mamta Juneja. Machine learning: A review of the algorithms and its applications. In Pradeep Kumar Singh, Arpan Kumar Kar, Yashwant Singh, Maheshkumar H. Kolekar, and Sudeep Tanwar, editors, *Proceedings of ICRIC 2019*, pages 47–63, Cham, 2020. Springer International Publishing.

181. Larry V. Hedges. Research synthesis: The state of the art. *The International Journal of Aging and Human Development*, 19(2):85–93, 1984. doi: 10.

2190/VKWU-6PMU-LYJD-4BX6. URL https://doi.org/10.2190/VKWU-6PMU-LYJD-4BX6. PMID: 6519822.

182. "Rahul C. Deo". "machine learning in medicine". *"Circulation"*, "132"("20"): "1920–1930, 2015. doi: 10.1161/CIRCULATIONAHA.115.001593. URL https://www.ahajournals.org/doi/abs/10.1161/CIRCULATIONAHA.115.001593.

183. Mir Henglin, Gillian Stein, Pavel V. Hushcha, Jasper Snoek, Alexander B. Wiltschko, and Susan Cheng. Machine learning approaches in cardiovascular imaging. *Circulation: Cardiovascular Imaging*, 10(10):e005614, 2017. doi: 10.1161/CIRCIMAGING.117.005614. URL https://www.ahajournals.org/doi/abs/10.1161/CIRCIMAGING.117.005614.

184. G. S. Handelman, H. K. Kok, R. V. Chandra, A. H. Razavi, M. J. Lee, and H. Asadi. edoctor: machine learning and the future of medicine. *Journal of Internal Medicine*, 0(0), 08 2018. doi: 10.1111/joim.12822. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/joim.12822.

185. T Mitchell, B Buchanan, G DeJong, T Dietterich, P Rosenbloom, and A Waibel. Machine learning. *Annual Review of Computer Science*, 4(1):417–433, 1990. doi: 10.1146/annurev.cs.04.060190.002221. URL https://doi.org/10.1146/annurev.cs.04.060190.002221.

186. Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

187. Taiwo Oladipupo Ayodele. Types of machine learning algorithms. In Yagang Zhang, editor, *New Advances in Machine Learning*, chapter 3. IntechOpen, Rijeka, 2010. doi: 10.5772/9385. URL https://doi.org/10.5772/9385.

188. M. A. Alsheikh, S. Lin, D. Niyato, and H. Tan. Machine learning in wireless sensor networks: Algorithms, strategies, and applications. *IEEE Communications*

Surveys Tutorials, 16(4):1996–2018, Fourthquarter 2014. ISSN 1553-877X. doi: 10.1109/COMST.2014.2320099.

189. Sebastian Ruder. Transfer learning - machine learning's next frontier. *ruder.io*, 21 march 2017. URL `http://ruder.io/transfer-learning/`. post online. Retrieved on 2019/02/01.

190. Khader Shameer, Kipp W Johnson, Benjamin S Glicksberg, Joel T Dudley, and Partho P Sengupta. Machine learning in cardiovascular medicine: are we there yet? *Heart*, 104(14):1156–1164, 01 2018. ISSN 1355-6037. doi: 10.1136/heartjnl-2017-311198. URL `https://heart.bmj.com/content/104/14/1156`.

191. Rahul Kumar Sevakula, Wan-Tai M Au-Yeung, Jagmeet P Singh, E Kevin Heist, Eric M Isselbacher, and Antonis A Armoundas. State-of-the-art machine learning techniques aiming to improve patient outcomes pertaining to the cardiovascular system. *Journal of the American Heart Association*, 9(4):e013924, 2020. doi: 10.1161/JAHA.119.013924. URL `https://www.ahajournals.org/doi/abs/10.1161/JAHA.119.013924`.

192. George A. F. Seber and Alan J. Lee. *Linear Regression Analysis*. John Wiley & Sons, Ltd, seconde edition edition, Feb 2003. ISBN 9780471415404.

193. David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied logistic regression*, volume 398. John Wiley & Sons, Ltd, third edition edition, 04 2013. ISBN 9780470582473.

194. Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

195. Robert Tibshirani. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(3):273–282, 2011.

196. Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

197. Donald W Marquardt and Ronald D Snee. Ridge regression in practice. *The American Statistician*, 29(1):3–20, 1975.

198. Arthur E Hoerl, Robert W Kannard, and Kent F Baldwin. Ridge regression: some simulations. *Communications in Statistics-Theory and Methods*, 4(2):105–123, 1975.

199. Joseph O Ogutu, Torben Schulz-Streeck, and Hans-Peter Piepho. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. In *BMC proceedings*, volume 6, page S10. BioMed Central, 05 2012. doi: 10.1186/1753-6561-6-S2-S10. URL `https://pubmed.ncbi.nlm.nih.gov/22640436`.

200. Hui Zou and Trevor Hastie. Regression shrinkage and selection via the elastic net, with applications to microarrays. *JR Stat Soc Ser B*, 67:301–20, 2003.

201. Chris Hans. Elastic net regression modeling with the orthant normal prior. *Journal of the American Statistical Association*, 106(496):1383–1393, 2011.

202. Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67 (2):301–320, 2005.

203. Edward I Altman, Giancarlo Marco, and Franco Varetto. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the italian experience). *Journal of banking & finance*, 18(3):505–529, 1994.

204. Ming Li and Baozong Yuan. 2D-LDA: A statistical linear discriminant analysis for image matrix. *Pattern Recognition Letters*, 26(5):527–532, 2005.

205. Suresh Balakrishnama and Aravind Ganapathiraju. Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing*, 18:1–8, 1998.

206. Ron Kohavi. Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In *Kdd*, volume 96, pages 202–207, 1996.

207. Kevin P Murphy et al. Naive bayes classifiers. *University of British Columbia*, 18: 60, 2006.

208. Fuchun Peng, Dale Schuurmans, and Shaojun Wang. Augmenting naive bayes classifiers with statistical language models. *Information Retrieval*, 7(3-4):317–345, 2004.

209. Dewan Md Farid, Li Zhang, Chowdhury Mofizur Rahman, M Alamgir Hossain, and Rebecca Strachan. Hybrid decision tree and naïve bayes classifiers for multi-class classification tasks. *Expert systems with applications*, 41(4):1937–1946, 2014.

210. Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, COLT'92, pages 144–152, New York, NY, USA, 1992. Association for Computing Machinery. ISBN 089791497X. doi: 10.1145/130385.130401. URL `https://doi.org/10.1145/130385.130401`.

211. Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.

212. Gert Cauwenberghs and Tomaso Poggio. Incremental and decremental support vector machine learning. In *Advances in neural information processing systems*, pages 409–415, 2001.

213. Achmad Widodo and Bo-Suk Yang. Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical systems and signal processing*, 21(6): 2560–2574, 2007.

214. William S Noble. What is a support vector machine? *Nature biotechnology*, 24 (12):1565–1567, 2006.

215. Thomas Dyhre Nielsen and Finn Verner Jensen. *Bayesian networks and decision graphs*. Springer Science & Business Media, 2009.

216. Finn V Jensen et al. *An introduction to Bayesian networks*, volume 210. UCL press London, 1996.

217. Uffe B Kjaerulff and Anders L Madsen. Bayesian networks and influence diagrams. *Springer Science+ Business Media*, 200:114, 2008.

218. J. Ross Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.

219. Lior Rokach and Oded Z Maimon. *Data mining with decision trees: theory and applications*, volume 69. World scientific, 2008.

220. Helmut Schmid. Probabilistic part-ofispeech tagging using decision trees. In *New methods in language processing*, page 154, 2013.

221. Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

222. Andy Liaw, Matthew Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.

223. Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, 28(2):337–407, 2000.

224. Peter L Bartlett and Mikhail Traskin. Adaboost is consistent. *Journal of Machine Learning Research*, 8(Oct):2347–2368, 2007.

225. Mohamad H Hassoun et al. *Fundamentals of artificial neural networks*. MIT press, 1995.

226. Xin Yao. Evolving artificial neural networks. *Proceedings of the IEEE*, 87(9): 1423–1447, 1999.

227. Guoqiang Zhang, B Eddy Patuwo, and Michael Y Hu. Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, 14(1): 35–62, 1998.

228. Imad A Basheer and Maha Hajmeer. Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, 43(1): 3–31, 2000.

229. Graupe Daniel. *Principles of artificial neural networks*, volume 7. World Scientific, 2013.

230. Bayya Yegnanarayana. *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.

231. Zhi-Hua Zhou. *Ensemble methods: foundations and algorithms*. CRC press, 2012.

232. Thomas G Dietterich. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, pages 1–15. Springer, 2000.

233. Zhi-Hua Zhou. Ensemble learning. *Encyclopedia of biometrics*, 1:270–273, 2009.

234. Roger N Anderson, Albert Boulanger, Leon L Wu, Viabhav Bhandari, Somnath Sarkar, and Ashish Gagneja. Forecasting system using machine learning and ensemble methods, November 5 2015. US Patent App. 14/707,809.

235. Przemysław Kazienko, Edwin Lughofer, and Bogdan Trawiński. Hybrid and ensemble methods in machine learning j. ucs special issue. *J Univers Comput Sci*, 19(4):457–461, 2013.

236. L. S. Mihalkova. *Learning with Markov Logic Networks: Transfer Learning, Structure Learning, and an Application to Web Query Disambiguation*. University of Texas, 2009. URL `https://books.google.fr/books?id=uOIunQAACAAJ`.

237. Waseem Rawat and Zenghui Wang. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 29(9):2352–2449, 2017. doi: 10.1162/neco_a_00990. URL https://doi.org/10.1162/neco_a_00990. PMID: 28599112.

238. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014. URL http://arxiv.org/abs/1409.1556.

239. Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567, 2015. URL http://arxiv.org/abs/1512.00567.

240. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. URL http://arxiv.org/abs/1512.03385.

241. Alfredo Canziani, Adam Paszke, and Eugenio Culurciello. An analysis of deep neural network models for practical applications. *CoRR*, abs/1605.07678, 2016. URL http://arxiv.org/abs/1605.07678.

242. Manu Sharma, Michael Holmes, Juan Santamaria, Arya Irani, Charles Isbell, and Ashwin Ram. Transfer Learning in Real-time Strategy Games Using Hybrid CBR/RL. In *Proceedings of the 20th International Joint Conference on Artifical Intelligence*, IJCAI'07, pages 1041–1046, San Francisco, CA, USA, 2007. Morgan Kaufmann Publishers Inc. URL http://dl.acm.org/citation.cfm?id=1625275.1625444.

243. S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, Oct 2010. ISSN 1041-4347. doi: 10.1109/TKDE.2009.191.

244. Qiang Yang and Xindong Wu. 10 challenging problems in data mining research. *International Journal of Information Technology & Decision Making*, 05(04):597–604, 2006. doi: 10.1142/S0219622006002258. URL https://doi.org/10.1142/S0219622006002258.

245. L. Torrey and J. Shavlik. Transfer learning. In E. Soria, J. Martin, R. Magdalena, M. Martinez, and A. Serrano, editors, *Handbook of Research on Machine Learning Applications*. IGI Global, 2009.

246. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. URL http://www.deeplearningbook.org.

247. Jeremy West, Dan Ventura, and Sean Warnick. Spring research presentation: A theoretical foundation for inductive transfer (abstract only). *Brigham Young University, College of Physical and Mathematical Sciences.*, 2007. URL https://web.archive.org/web/20070801120743/http://cpms.byu.edu/springresearch/abstract-entry?id=861. Archived from the original on 2007-08-01. Retrieved on 2019-01-23.

248. Emilio Soria Olivas, Jose David Martin Guerrero, Marcelino Martinez Sober, Jose Rafael Magdalena Benedito, and Antonio Jose Serrano Lopez. *Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques - 2 Volumes*. Information Science Reference - Imprint of: IGI Publishing, Hershey, PA, 2009. ISBN 1605667668, 9781605667669.

249. Athanasios Voulodimos, Nikolaos Doulamis, Anastasios Doulamis, and Eftychios Protopapadakis. Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience*, 2018, 2018. doi: 10.1155/2018/7068349. URL https://doi.org/10.1155/2018/7068349. Article ID 7068349.

250. D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591, 1959. doi: 10.1113/jphysiol.

1959.sp006308. URL https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1959.sp006308.

251. D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962. doi: 10.1113/jphysiol.1962.sp006837. URL https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1962.sp006837.

252. Ayon Dey. Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies*, 7(3):1174–1179, 2016.

253. BYJU'S. The Learning App. Neuron and nerves, 2020. URL https://byjus.com/biology/neuron-and-nerves/. [Online; accessed 3-May-2020].

254. Vidushi Sharma, Sachin Rai, and Anurag Dev. A comprehensive study of artificial neural networks. *International Journal of Advanced research in computer science and software engineering*, 2(10), 2012.

255. Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, pages 807–814, USA, 2010. Omnipress. ISBN 978-1-60558-907-7. URL http://dl.acm.org/citation.cfm?id=3104322.3104425.

256. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'12, pages 1097–1105, USA, 2012. Curran Associates Inc. URL http://dl.acm.org/citation.cfm?id=2999134.2999257.

257. Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. *CoRR*, abs/1312.4400, 2013. URL http://arxiv.org/abs/1312.4400.

258. Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 818–833, Cham, 2014. Springer International Publishing. ISBN 978-3-319-10590-1.

259. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott E. Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014. URL http://arxiv.org/abs/1409.4842.

260. Bing Xu, Naiyan Wang, Tianqi Chen, and Mu Li. Empirical evaluation of rectified activations in convolutional network. *CoRR*, abs/1505.00853, 2015. URL http://arxiv.org/abs/1505.00853.

261. Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521:436–444, 2015. doi: 10.1038/nature14539. URL https://doi.org/10.1038/nature14539.

262. Dishashree Gupta. 25 Must Know Terms & concepts for Beginners in Deep Learning, 05 2017. URL https://www.analyticsvidhya.com/blog/2017/05/25-must-know-terms-concepts-for-beginners-in-deep-learning/. [Online post; accessed 24-April-2020].

263. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4):541–551, Dec 1989. ISSN 0899-7667. doi: 10.1162/neco.1989.1.4.541.

264. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, Nov 1998. ISSN 0018-9219. doi: 10.1109/5.726791.

265. Yoshua Bengio. Learning Deep Architectures for AI. *Foundations and Trends in Machine Learning*, 2(1):1–127, 2009. ISSN 1935-8237. doi: 10.1561/2200000006. URL http://dx.doi.org/10.1561/2200000006.

266. Li Deng and Dong Yu. Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, 7(3–4):197–387, 2014. ISSN 1932-8346. doi: 10.1561/2000000039. URL http://dx.doi.org/10.1561/2000000039.

267. Li Deng. A tutorial survey of architectures, algorithms, and applications for deep learning. *APSIPA Transactions on Signal and Information Processing*, January 2014. URL https://www.microsoft.com/en-us/research/publication/a-tutorial-survey-of-architectures-algorithms-and-applications-for-deep-learning/.

268. Suraj Srinivas, Ravi Kiran Sarvadevabhatla, Konda Reddy Mopuri, Nikita Prabhu, Srinivas S. S. Kruthiventi, and R. Venkatesh Babu. A taxonomy of deep convolutional neural nets for computer vision. *CoRR*, abs/1601.06615, 2016. URL http://arxiv.org/abs/1601.06615.

269. I. Aizenberg, N. Aizenberg, J. Hiltner, C. Moraga, and E. Meyer zu Bexten. Cellular neural networks and computational intelligence in medical image processing. *Image and Vision Computing*, 19(4):177 – 183, 2001. ISSN 0262-8856. doi: https://doi.org/10.1016/S0262-8856(00)00066-4. URL http://www.sciencedirect.com/science/article/pii/S0262885600000664.

270. M. Egmont-Petersen, D. de Ridder, and H. Handels. Image processing with neural networks–a review. *Pattern Recognition*, 35(10):2279 – 2301, 2002. ISSN 0031-3203. doi: https://doi.org/10.1016/S0031-3203(01)00178-9. URL http://www.sciencedirect.com/science/article/pii/S0031320301001789.

271. François Chollet. *Deep learning with Python*. Manning Publications Co., 20

Baldwin Road PO Bos 761, Shelter Island, NY 11964, USA, 2018. ISBN 9781617294433.

272. Yann Lecun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L.D. Jackel. Handwritten digit recognition with a back-propagation network. In David Touretzky, editor, *Advances in Neural Information Processing Systems (NIPS 1989), Denver, CO*, volume 2. Morgan Kaufmann, 1990.

273. M. Ranzato, F. J. Huang, Y. Boureau, and Y. LeCun. Unsupervised learning of invariant feature hierarchies with applications to object recognition. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007. doi: 10.1109/CVPR.2007.383157.

274. Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *CoRR*, abs/1207.0580, 2012. URL `http://arxiv.org/abs/1207.0580`.

275. Shiva Ebrahimi Nejad, Jason P Carey, M Sean McMurtry, and Jin-Oh Hahn. Model-based cardiovascular disease diagnosis: a preliminary in-silico study. *Biomechanics and modeling in mechanobiology*, 16(2):549–560, 2017.

276. Shanthi Mendis, Pekka Puska, Bo Norrving, World Health Organization, et al. *Global atlas on cardiovascular disease prevention and control*. Geneva: World Health Organization, 2011.

277. Robert Rapadamnaba, Franck Nicoud, and Bijan Mohammadi. Backward sensitivity analysis and reduced-order covariance estimation in noninvasive parameter identification for cerebral arteries. *International Journal for Numerical Methods in Biomedical Engineering*, 35(4):e3170, 2019. doi: 10.1002/cnm.3170. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3170`. e3170 cnm.3170.

278. Robert Rapadamnaba, Franck Nicoud, and Bijan Mohammadi. Augmented patient-specific functional medical imaging by implicit manifold learning. *International Journal for Numerical Methods in Biomedical Engineering*, 36(5): e3325, 2020. doi: 10.1002/cnm.3325. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3325.

279. World Health Organization. Cardiovascular diseases (CVDs). The United Nations, 05 2017. URL <http://www.who.int/mediacentre/factsheets/fs317/en/>. [Online; accessed 28-March-2019].

280. Brain Aneurysm Foundation. About brain aneurysms, 2018. URL https://www.bafound.org/about-brain-aneurysms/brain-aneurysm-basics/warning-signs-symptoms/.

281. M. Sanchez, O. Ecker, D. Ambard, F. Jourdan, F. Nicoud, S. Mendez, J.-P. Lejeune, L. Thines, H. Dufour, H. Brunel, P. Machi, K. Lobotesis, A. Bonafe, and V. Costalat. Intracranial aneurysmal pulsatility as a new individual criterion for rupture risk evaluation: Biomechanical and numeric approach (irras project). *American Journal of Neuroradiology*, 35(9):1765–1771, 2014. ISSN 0195-6108. doi: 10.3174/ajnr.A3949. URL http://www.ajnr.org/content/35/9/1765.

282. Nima Etminan, Bruce A. Buchholz, Rita Dreier, Peter Bruckner, James C. Torner, Hans-Jakob Steiger, Daniel Hänggi, and Macdonald R. Loch. Cerebral aneurysms: Formation, progression and developmental chronology. *Translational stroke research*, 2(5):167–173, 2014. doi: 10.1007/s12975-013-0294-x.

283. Masaaki Shojima, Marie Oshima, Kiyoshi Takagi, Ryo Torii, Motoharu Hayakawa, Kazuhiro Katada, Akio Morita, and Takaaki Kirino. Magnitude and role of wall shear stress on cerebral aneurysm. *Stroke*, 35(11):2500–2505, 2004. doi: 10.1161/01.STR.0000144648.89172.0f. URL https://www.ahajournals.org/doi/abs/10.1161/01.STR.0000144648.89172.0f.

284. J Alastruey, KH Parker, J Peiró, SM Byrd, and SJ Sherwin. Modelling the circle of willis to assess the effects of anatomical variations and occlusions on cerebral flows. *Journal of biomechanics*, 40(8):1794–1805, 2007.

285. K N Kayembe, M Sasahara, and F Hazama. Cerebral aneurysms and variations in the circle of willis. *American Heart Association, Inc.*, 15(5):846–850, 1984. ISSN 0039-2499. doi: 10.1161/01.STR.15.5.846. URL http://stroke.ahajournals.org/content/15/5/846.

286. Philippe Reymond, Fabrice Merenda, Fabienne Perren, Daniel Rüfenacht, and Nikos Stergiopulos. Validation of a one-dimensional model of the systemic arterial tree. *American Journal of Physiology-Heart and Circulatory Physiology*, 297(1): H208–H222, 2009.

287. G Mulder, ACB Bogaerds, P Rongen, and FN van de Vosse. The influence of contrast agent injection on physiological flow in the circle of willis. *Medical engineering & physics*, 33(2):195–203, 2011.

288. Ryo Torii, Marie Oshima, Toshio Kobayashi, Kiyoshi Takagi, and Tayfun E. Tezduyar. Fluid–structure interaction modeling of blood flow and cerebral aneurysm: Significance of artery and aneurysm shapes. *Computer Methods in Applied Mechanics and Engineering*, 198(45):3613 – 3621, 2009. ISSN 0045-7825. doi: https://doi.org/10.1016/j.cma.2008.08.020. URL http://www.sciencedirect.com/science/article/pii/S0045782508003174. Models and Methods in Computational Vascular and Cardiovascular Mechanics.

289. Qingzhuo Chi, Xinyu Li, Sheng Chang, Lizhong Mu, Ying He, and Ge Gao. In-vitro experimental study on the fluid-structure interaction in an image-based flexible model with a lateral cerebral aneurysm. *Medicine in Novel Technology and Devices*, 3:100019, 2019. ISSN 2590-0935. doi: https://doi.org/10.1016/j.medntd.2019.

100019. URL http://www.sciencedirect.com/science/article/pii/S2590093519300190.

290. Ryo Torii, Marie Oshima, Toshio Kobayashi, Kiyoshi Takagi, and Tayfun E. Tezduyar. Coupling 3d fluid-structure interaction modeling of cerebral aneurysm with 0d arterial network model as boundary conditions. *Transaction of the Japan Society for Simulation Technology*, 1(4):81–90, 2009. doi: 10.11308/tjsst.1.81.

291. Ryo Torii, Marie Oshima, Toshio Kobayashi, Kiyoshi Takagi, and Tayfun E. Tezduyar. Role of 0d peripheral vasculature model in fluid-structure interaction modeling of aneurysms. *Computational Mechanics*, 46(1):43–52, June 2010. ISSN 0178-7675. doi: 10.1007/s00466-009-0439-7.

292. Christof Karmonik, Christopher Yen, Robert G. Grossman, Richard Klucznik, and Goetz Benndorf. Intra-aneurysmal flow patterns and wall shear stresses calculated with computational flow dynamics in an anterior communicating artery aneurysm depend on knowledge of patient-specific inflow rates. *Acta Neurochirurgica*, 151 (5):479–485, May 2009. ISSN 0942-0940. doi: 10.1007/s00701-009-0247-z. URL https://doi.org/10.1007/s00701-009-0247-z.

293. Tamer Hassan, Amr Ali Hassan, and Yaser Mohamed Ahmed. Influence of parent vessel dominancy on fluid dynamics of anterior communicating artery aneurysms. *Acta Neurochirurgica*, 153(2):305–310, Feb 2011. ISSN 0942-0940. doi: 10.1007/s00701-010-0824-1. URL https://doi.org/10.1007/s00701-010-0824-1.

294. Marc A Lazzaro, Bichun Ouyang, and Michael Chen. The role of circle of willis anomalies in cerebral aneurysm rupture. *Journal of NeuroInterventional Surgery*, 4(1):22–26, 2012. ISSN 1759-8478. doi: 10.1136/jnis.2010.004358. URL https://jnis.bmj.com/content/4/1/22.

295. Huang P. G. and Muller L. O. Simulation of one-dimensional blood flow in networks of human vessels using a novel tvd scheme. *International Journal for Numerical Methods in Biomedical Engineering*, 31(5):e02701, 2015. doi: 10.1002/cnm.2701. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2701. e02701 CNM-Sep-14-0153.R1.

296. Bruno C. Flores, William W. Scott, Christopher S. Eddleman, H. Hunt Batjer, and Kim L. Rickert. The a1-a2 diameter ratio may influence formation and rupture potential of anterior communicating artery aneurysms. *Neurosurgery*, 73(5):845–853, 2013. doi: 10.1227/NEU.0000000000000125. URL http://dx.doi.org/10.1227/NEU.0000000000000125.

297. Fuyou Liang, Xiaosheng Liu, Ryuhei Yamaguchi, and Hao Liu. Sensitivity of flow patterns in aneurysms on the anterior communicating artery to anatomic variations of the cerebral arterial network. *Journal of Biomechanics*, 49(15):3731–3740, Nov 2016. doi: doi:10.1016/j.jbiomech.2016.09.031. URL https://doi.org/10.1016/j.jbiomech.2016.09.031.

298. Ryo Torii, Marie Oshima, Toshio Kobayashi, Kiyoshi Takagi, and Tayfun E. Tezduyar. Influencing factors in image-based fluid–structure interaction computation of cerebral aneurysms. *International Journal for Numerical Methods in Fluids*, 65(1-3):324–340, 2011. doi: 10.1002/fld.2448. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/fld.2448.

299. Cristóbal Bertoglio, Philippe Moireau, and Jean-Frederic Gerbeau. Sequential parameter estimation for fluid–structure problems: Application to hemodynamics. *International Journal for Numerical Methods in Biomedical Engineering*, 28(4):434–455, 2012.

300. Philippe Moireau, Cristobal Bertoglio, Nan Xiao, C Alberto Figueroa, CA Taylor, Dominique Chapelle, and J-F Gerbeau. Sequential identification of boundary

support parameters in a fluid-structure vascular model using patient image data. *Biomechanics and modeling in mechanobiology*, 12(3):475–496, 2013.

301. Damiano Lombardi. Inverse problems in 1D hemodynamics on systemic networks: A sequential approach. *International journal for numerical methods in biomedical engineering*, 30(2):160–179, 2014.

302. Radomir Chabiniok, Philippe Moireau, P-F Lesault, Alain Rahmouni, J-F Deux, and Dominique Chapelle. Estimation of tissue contractility from cardiac cine-mri using a biomechanical heart model. *Biomechanics and modeling in mechanobiology*, 11(5):609–630, 2012.

303. A Caiazzo, F Caforio, G Montecinos, LO Muller, PJ Blanco, and EF Toro. Assessment of reduced-order unscented kalman filter for parameter identification in one-dimensional blood flow models using experimental data. *International Journal for Numerical Methods in Biomedical Engineering*, 2016.

304. Benoit Fabrèges, Sanjay Pant, Irene Vignon-Clementel, and Jean-Frédéric Gerbeau. Parameter estimation for a 3D Navier-Stokes - 0D coupled system: application to patient-specific haemodynamics. In *3rd International Conference on Computational & Mathematical Biomedical Engineering*, Hong Kong, China, December 2013. URL `https://hal.inria.fr/hal-00918373`.

305. Sanjay Pant, Benoit Fabrèges, J-F Gerbeau, and IE Vignon-Clementel. A methodological paradigm for patient-specific multi-scale cfd simulations: from clinical measurements to parameter estimates for individual analysis. *International journal for numerical methods in biomedical engineering*, 30(12):1614–1648, 2014.

306. Sanjay Pant, Chiara Corsini, Catriona Baker, Tain-Yen Hsia, Giancarlo Pennati, Irene E Vignon-Clementel, Modeling of Congenital Hearts Alliance (MOCHA) Investigators, et al. Data assimilation and modelling of patient-

specific single-ventricle physiology with and without valve regurgitation. *Journal of biomechanics*, 2015.

307. Sanjay Pant, Chiara Corsini, Catriona Baker, Tain-Yen Hsia, Giancarlo Pennati, Irene E Vignon-Clementel, Modeling of Congenital Hearts Alliance (MOCHA) Investigators, et al. Data assimilation and modelling of patient-specific single-ventricle physiology with and without valve regurgitation. *Journal of biomechanics*, 49(11):2162–2173, 2016.

308. Sanjay Pant, Chiara Corsini, Catriona Baker, Tain-Yen Hsia, Giancarlo Pennati, and Irene E Vignon-Clementel. Inverse problems in reduced order models of cardiovascular haemodynamics: aspects of data assimilation and heart rate variability. *Journal of The Royal Society Interface*, 14(126):20160513, 2017.

309. Sanjay Pant, Benoit Fabrèges, Jean-Frédéric Gerbeau, and Irene Vignon-Clementel. A multiscale filtering-based parameter estimation method for patient-specific coarctation simulations in rest and exercise. In *STACOM, MICCAI - 4th International Workshop on Statistical Atlases and Computational Models of the Heart held in conjunction with the 16th International conference on Medical Image Computing and Computer Assisted Intervention, 2013*, Nagoya, Japan, September 2013. URL `https://hal.inria.fr/hal-00911339`.

310. B. Mohammadi. Backward uncertainty propagation in shape optimization. *Int. J. for Numerical Methods in Fluids*, 103(4)-DOI: 10.1002/fld.4077:307–323, 2016.

311. B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids (2nd Edition)*. Oxford Univ. Press, Oxford, 2009.

312. Alfio Quarteroni, Massimiliano Tuveri, and Alessandro Veneziani. Computational vascular fluid dynamics: problems, models and methods. *Computing and Visualization in Science*, 2(4):163–97, 2000. doi: 10.1007/s007910050039.

313. I.N. Kiselev, Boris Semisalov, Ruslan Sharipov, and Fedor Kolpakov. *Modular modeling of the human cardiovascular system*, volume 7. Nova Science Publishers Inc., 11 2012. ISBN 978-1-62257-275-5.

314. Fuyou Liang, Shu Takagi, Ryutaro Himeno, and Hao Liu. Multi-scale modeling of the human cardiovascular system with applications to aortic valvular and arterial stenoses. *Medical & Biological Engineering & Computing*, 47(7):743–755, Jul 2009. ISSN 1741-0444. doi: 10.1007/s11517-009-0449-9. URL https://doi.org/10.1007/s11517-009-0449-9.

315. Mauro Ursino and Massimo Giannessi. A model of cerebrovascular reactivity including the circle of willis and cortical anastomoses. *Annals of biomedical engineering*, 38(3):955–974, 2010.

316. Alfio Quarteroni, Stefania Ragni, and Alessandro Veneziani. Coupling between lumped and distributed models for blood flow problems. *Computing and Visualization in Science*, 4(2):111–124, 2001.

317. Mette S Olufsen, Ali Nadim, et al. On deriving lumped models for blood flow and pressure in the systemic arteries. *Math Biosci Eng*, 1(1):61–80, 2004.

318. Etienne Boileau, Perumal Nithiarasu, Pablo J. Blanco, Lucas O. Müller, Fredrik Eikeland Fossan, Leif Rune Hellevik, Wouter P. Donders, Wouter Huberts, Marie Willemet, and Jordi Alastruey. A benchmark study of numerical schemes for one-dimensional arterial blood flow modelling. *International Journal for Numerical Methods in Biomedical Engineering*, 31(10):e02732, oct 2015. doi: 10.1002/cnm.2732. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2732. e02732 cnm.2732.

319. George D Byrne and Alan C Hindmarsh. Stiff ODE solvers: A review of current and coming attractions. *Journal of Computational physics*, 70(1):1–62, 1987.

320. S Thompson SP Corwin and SM White. Solving ODEs and DDEs with impulses. *JNAIAM J. Numer. Anal .Indust. Appl. Math*, 3:139–149, 2008.

321. Mette S Olufsen. Structured tree outflow condition for blood flow in larger systemic arteries. *American journal of physiology-Heart and circulatory physiology*, 276(1): H257–H268, 1999.

322. R. Gul, C. Schütte, and S. Bernhard. Mathematical modeling and sensitivity analysis of arterial anastomosis in the arm. *Applied Mathematical Modelling*, 40 (17):7724 – 7738, 2016. ISSN 0307-904X. doi: https://doi.org/10.1016/j.apm.2016.03.041. URL http://www.sciencedirect.com/science/article/pii/S0307904X16301895.

323. Fuyou Liang, Kazuaki Fukasaku, Hao Liu, and Shu Takagi. A computational model study of the influence of the anatomy of the circle of willis on cerebral hyperperfusion following carotid artery surgery. *Biomedical engineering online*, 10 (1):84, 2011.

324. Jonas Schollenberger. A lumped parameter model of cerebral blood flow regulation: Applications to simulation of carotid endarterectomy. *Technische Universitat Munchen, Master Thesis*, 06 2015. URL https://bloodflow.engin.umich.edu/wp-content/uploads/sites/165/2014/07/20141201_thesis.pdf.

325. L. Wasserman. *All of Statistics: A Concise Course in Statistical Inference*. ISBN 0-387-40272-1. Springer, 2004.

326. Nico Westerhof, Jan-Willem Lankhaar, and Berend E. Westerhof. The arterial windkessel. *Medical & Biological Engineering & Computing*, 47(2):131–141, Feb 2009. ISSN 1741-0444. doi: 10.1007/s11517-008-0359-2. URL https://doi.org/10.1007/s11517-008-0359-2.

327. Luca Formaggia, Jean-Frederic Gerbeau, Fabio Nobile, and Alfio Quarteroni. On the Coupling of 3D and 1D Navier-Stokes Equations for Flow Problems in Compliant Vessels. *Computer Methods in Applied Mechanics and Engineering*, 191:561–582, 12 2001.

328. Irene E. Vignon-Clementel, C. Alberto Figueroa, Kenneth E. Jansen, and Charles A. Taylor. Outflow boundary conditions for three-dimensional finite element modeling of blood flow and pressure in arteries. *Computer Methods in Applied Mechanics and Engineering*, 195(29):3776 – 3796, 2006. ISSN 0045-7825. doi: https://doi.org/10.1016/j.cma.2005.04.014. URL http://www.sciencedirect.com/science/article/pii/S0045782505002586. Absorbing Boundary Conditions.

329. H. J. Kim, I. E. Vignon-Clementel, C. A. Figueroa, J. F. LaDisa, K. E. Jansen, J. A. Feinstein, and C. A. Taylor. On coupling a lumped parameter heart model and a three-dimensional finite element aorta model. *Annals of Biomedical Engineering*, 37(11):2153–2169, Nov 2009. ISSN 1573-9686. doi: 10.1007/s10439-009-9760-8. URL https://doi.org/10.1007/s10439-009-9760-8.

330. Irene E. Vignon and Charles A. Taylor. Outflow boundary conditions for one-dimensional finite element modeling of blood flow and pressure waves in arteries. *Wave Motion*, 39(4):361 – 374, 2004. ISSN 0165-2125. doi: https://doi.org/10.1016/j.wavemoti.2003.12.009. URL http://www.sciencedirect.com/science/article/pii/S0165212503001173. New computational methods for wave propagation.

331. Hermann Schichl. *Models and the History of Modeling*, volume 88, chapter 2, pages 25–36. Springer US, Boston, MA, 1 edition, 2004. ISBN 978-1-4613-0215-5. doi: 10.1007/978-1-4613-0215-5_2. URL https://doi.org/10.1007/978-1-4613-0215-5_2. (J. Kallrath, ed.).

332. Arnold Neumaier. *Mathematical Model Building*, volume 88, chapter 3, pages 37–43. Springer US, Boston, MA, 1 edition, 2004. ISBN 978-1-4613-0215-5. doi: 10.1007/978-1-4613-0215-5_3. URL https://doi.org/10.1007/978-1-4613-0215-5_3. (J. Kallrath, ed.).

333. Michele Crosetto, Stefano Tarantola, and Andrea Saltelli. Sensitivity and uncertainty analysis in spatial modelling based on gis. *Agriculture, Ecosystems & Environment*, 81(1):71 – 79, 2000. ISSN 0167-8809. doi: https://doi.org/10.1016/S0167-8809(00)00169-9. URL http://www.sciencedirect.com/science/article/pii/S0167880900001699.

334. David J. Pannell. Sensitivity analysis of normative economic models: theoretical framework and practical strategies. *Agricultural Economics*, 16(2):139 – 152, 1997. ISSN 0169-5150. doi: https://doi.org/10.1016/S0169-5150(96)01217-0. URL http://www.sciencedirect.com/science/article/pii/S0169515096012170.

335. Mary C. Hill, Dmitri Kavetski, Martyn Clark, Ming Ye, Mazdak Arabi, Dan Lu, Laura Foglia, and Steffen Mehl. Practical use of computationally frugal model analysis methods. *Groundwater*, 54(2):159–170, 2016. doi: 10.1111/gwat.12330. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/gwat.12330.

336. Francesca Pianosi, Keith Beven, Jim Freer, Jim W. Hall, Jonathan Rougier, David B. Stephenson, and Thorsten Wagener. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79:214 – 232, 2016. ISSN 1364-8152. doi: https://doi.org/10.1016/j.envsoft.2016.02.008. URL http://www.sciencedirect.com/science/article/pii/S1364815216300287.

337. T Lenhart, K Eckhardt, N Fohrer, and H.-G Frede. Comparison of two different approaches of sensitivity analysis. *Physics and Chemistry of the Earth, Parts*

A/B/C, 27(9):645 – 654, 2002. ISSN 1474-7065. doi: https://doi.org/10.1016/S1474-7065(02)00049-9. URL http://www.sciencedirect.com/science/article/pii/S1474706502000499.

338. A. Bahremand and F. De Smedt. Distributed hydrological modeling and sensitivity analysis in torysa watershed, slovakia. *Water Resources Management*, 22(3):393–408, Mar 2008. ISSN 1573-1650. doi: 10.1007/s11269-007-9168-x. URL https://doi.org/10.1007/s11269-007-9168-x.

339. Mary C. Hill and Claire R. Tiedeman. *Effective Groundwater Model Calibration: With Analysis of Data, Sensitivities, Predictions, and Uncertainty*. Wiley, New Jersey, john wiley & sons, inc edition, 2007. ISBN 9780471776369. doi: 10.1002/0470041080.

340. Andrea Saltelli and E.Marian Scott. The role of sensitivity analysis in the corroboration of models and its links to model structural and parametric uncertainty (Guest editorial). *Reliability Engineering and System Safety*, 57:1–4, 1997. URL http://publications.jrc.ec.europa.eu/repository/handle/JRC14027.

341. A. Saltelli, S. Tarantola, and F. Campolongo. Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15(4):377–395, 2000. ISSN 08834237. URL http://www.jstor.org/stable/2676831.

342. Andrea Saltelli and Stefano Tarantola. On the relative importance of input factors in mathematical models. *Journal of the American Statistical Association*, 97(459):702–709, 2002. doi: 10.1198/016214502388618447. URL https://doi.org/10.1198/016214502388618447.

343. SJ Sherwin, V Franke, J Peiró, and K Parker. One-dimensional modelling of a vascular network in space-time variables. *Journal of Engineering Mathematics*, 47 (3-4):217–250, 2003.

344. B.S. Everitt. *Cambridge Dictionary of Statistics*. Cambridge University Press, 1998. ISBN 9780521593465. URL https://books.google.fr/books?id=0dByQgAACAAJ.

345. John Norton. An introduction to sensitivity assessment of simulation models. *Environmental Modelling & Software*, 69:166 – 174, 2015. ISSN 1364-8152. doi: https://doi.org/10.1016/j.envsoft.2015.03.020. URL http://www.sciencedirect.com/science/article/pii/S1364815215001085.

346. Francesca Pianosi, Fanny Sarrazin, and Thorsten Wagener. A matlab toolbox for global sensitivity analysis. *Environmental Modelling & Software*, 70:80 – 85, 2015. ISSN 1364-8152. doi: https://doi.org/10.1016/j.envsoft.2015.04.009. URL http://www.sciencedirect.com/science/article/pii/S1364815215001188.

347. Xiaomeng Song, Jianyun Zhang, Chesheng Zhan, Yunqing Xuan, Ming Ye, and Chonggang Xu. Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications. *Journal of Hydrology*, 523:739 – 757, 2015. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2015.02.013. URL http://www.sciencedirect.com/science/article/pii/S0022169415001249.

348. Bertrand Iooss and Andrea Saltelli. Introduction to sensitivity analysis. In Roger Ghanem, David Higdon, and Houman Owhadi, editors, *Handbook of uncertainty quantification*, pages 1103–1122, Switzerland, June 2017. Springer, Cham. ISBN 978-3-319-12385-1. doi: 10.1007/978-3-319-12385-1_31. URL https://doi.org/10.1007/978-3-319-12385-1_31.

349. Fanny Sarrazin, Francesca Pianosi, and Thorsten Wagener. Global sensitivity analysis of environmental models. *Environ. Model. Softw.*, 79(C):135–152, May 2016. ISSN 1364-8152. doi: 10.1016/j.envsoft.2016.02.005. URL https://doi.org/10.1016/j.envsoft.2016.02.005.

350. Kathryn van Werkhoven, Thorsten Wagener, Patrick Reed, and Yong Tang. Characterization of watershed model behavior across a hydroclimatic gradient. *Water Resources Research*, 44(1), 2008. doi: 10.1029/2007WR006271. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007WR006271.

351. Gürkan Sin, Krist V. Gernaey, Marc B. Neumann, Mark C.M. van Loosdrecht, and Willi Gujer. Global sensitivity analysis in wastewater treatment plant model applications: Prioritizing sources of uncertainty. *Water Research*, 45(2): 639 – 651, 2011. ISSN 0043-1354. doi: https://doi.org/10.1016/j.watres.2010.08.025. URL http://www.sciencedirect.com/science/article/pii/S0043135410005890.

352. A. Hartmann, T. Wagener, A. Rimmer, J. Lange, H. Brielmann, and M. Weiler. Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resources Research*, 49(6):3345–3358, 2013. doi: 10.1002/wrcr.20229. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20229.

353. R.C. Spear and G.M. Hornberger. Eutrophication in peel inlet—ii. identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, 14(1):43 – 49, 1980. ISSN 0043-1354. doi: https://doi.org/10.1016/0043-1354(80)90040-8. URL http://www.sciencedirect.com/science/article/pii/0043135480900408.

354. R. Singh, T. Wagener, R. Crane, M. E. Mann, and L. Ning. A vulnerability driven approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: Application to a watershed in Pennsylvania, USA. *Water Resources Research*, 50:3409–3427, April 2014. doi: 10.1002/2013WR014988.

355. Tamás Turányi. Sensitivity analysis of complex kinetic systems. tools and applications. *Journal of Mathematical Chemistry*, 5(3):203–248, Sep 1990. ISSN 1572-8897. doi: 10.1007/BF01166355. URL https://doi.org/10.1007/BF01166355.

356. Mary C. Hill and Claire R. Tiedeman. *Evaluating Estimated Parameter Values and Parameter Uncertainty*, chapter 7, pages 124–157. John Wiley & Sons, Ltd, 2005. ISBN 9780470041086. doi: 10.1002/9780470041086.ch7. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470041086.ch7.

357. Lennart Ljung. *System Identification*, pages 1–19. American Cancer Society, 2017. ISBN 9780471346081. doi: 10.1002/047134608X.W1046.pub2. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/047134608X.W1046.pub2.

358. Toshimitsu Homma and Andrea Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

359. Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964 – 979, 2008. ISSN 0951-8320. doi: https://doi.org/10.1016/j.ress.2007.04.002. URL http://www.sciencedirect.com/science/article/pii/S0951832007001329. Bayesian Networks in Dependability.

360. Xiaobo Zhou, Henry Lin, and Henry Lin. *Global Sensitivity Analysis*, pages 408–409. Springer US, Boston, MA, 1 edition, 2008. ISBN 978-0-387-35973-1. doi: 10.1007/978-0-387-35973-1_538. URL https://doi.org/10.1007/978-0-387-35973-1_538. (S. Shekhar and H. Xiong, eds.).

361. G. E. B. Archer, A. Saltelli, and I. M. Sobol. Sensitivity measures,anova-like techniques and the use of bootstrap. *Journal of Statistical Computation*

and Simulation, 58(2):99–120, 1997. doi: 10.1080/00949659708811825. URL https://doi.org/10.1080/00949659708811825.

362. I.M Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1):271 – 280, 2001. ISSN 0378-4754. doi: https://doi.org/10.1016/S0378-4754(00)00270-6. URL http://www.sciencedirect.com/science/article/pii/S0378475400002706. The Second IMACS Seminar on Monte Carlo Methods.

363. I.M. Sobol and S. Kucherenko. Derivative based global sensitivity measures and their link with global sensitivity indices. *Mathematics and Computers in Simulation*, 79/10:3009–3017, 2009.

364. Andrea Saltelli, Marco Ratto, Stefano Tarantola, and Francesca Campolongo. Sensitivity analysis practices: Strategies for model-based inference. *Reliability Engineering & System Safety*, 91(10):1109 – 1125, 2006. ISSN 0951-8320. doi: https://doi.org/10.1016/j.ress.2005.11.014. URL http://www.sciencedirect.com/science/article/pii/S095183200500222X. The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004).

365. Alessandro Melis, Richard H. Clayton, and Alberto Marzo. Bayesian sensitivity analysis of a 1D vascular model with Gaussian process emulators. *International Journal for Numerical Methods in Biomedical Engineering*, 33(12):e2882, 2017. doi: 10.1002/cnm.2882. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2882. e2882 cnm.2882.

366. I. M. Sobol. Multidimensional quadrature formulas and haar functions. *Izdat "Nauka", Moscow*, 1969. (in Russian).

367. Bertrand Iooss and Paul Lemaître. *A Review on Global Sensitivity Analysis Methods*, volume 59 of *Operations Research/Computer Science Interfaces Series*, chapter 5, pages 101–122. Springer US, Boston, MA, 1 edition, 2015. ISBN 978-1-4899-7547-8. doi: 10.1007/978-1-4899-7547-8_5. URL https://doi.org/10.1007/978-1-4899-7547-8_5. (G. Dellino and C. Meloni, eds).

368. Julien Jacques, Christian Lavergne, and Nicolas Devictor. Sensitivity analysis in presence of model uncertainty and correlated inputs. *Reliability Engineering & System Safety*, 91(10):1126–1134, 2006. ISSN 0951-8320. doi: https://doi.org/10.1016/j.ress.2005.11.047. URL http://www.sciencedirect.com/science/article/pii/S0951832005002231. The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004).

369. Andrea Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280 – 297, 2002. ISSN 0010-4655. doi: https://doi.org/10.1016/S0010-4655(02)00280-1. URL http://www.sciencedirect.com/science/article/pii/S0010465502002801.

370. M. D. McKay, R. J. Beckman, and W. J. Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979. doi: 10.1080/00401706.1979.10489755. URL https://doi.org/10.1080/00401706.1979.10489755.

371. IM Sobol'. Quasi-monte carlo methods. *Progress in Nuclear Energy*, 24(1-3):55–61, 1990.

372. Harald Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992. doi: 10.1137/1.9781611970081. URL https://epubs.siam.org/doi/abs/10.1137/1.9781611970081.

373. Art B. Owen. Randomly permuted (t,m,s)-nets and (t, s)-sequences. In Harald Niederreiter and Peter Jau-Shyong Shiue, editors, *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, volume 106 of *LectureNotes in Statistics*, pages 299–317, New York, NY, 1995. Springer New York. ISBN 978-1-4612-2552-2.

374. A. B. Owen. Monte carlo extension of quasi-monte carlo. In *1998 Winter Simulation Conference. Proceedings (Cat. No.98CH36274)*, volume 1, pages 571–577 vol.1, Dec 1998. doi: 10.1109/WSC.1998.745036.

375. Art B. Owen. Monte carlo, quasi-monte carlo, and randomized quasi-monte carlo. In Harald Niederreiter and Jerome Spanier, editors, *Monte-Carlo and Quasi-Monte Carlo Methods 1998*, pages 86–97, Berlin, Heidelberg, 01 2000. Springer Berlin Heidelberg. ISBN 9783642596575. doi: 10.1007/978-3-642-59657-5_5.

376. Andrea Saltelli and Paola Annoni. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12):1508–1517, 2010.

377. RI Cukier, HB Levine, and KE Shuler. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of computational physics*, 26(1):1–42, 1978.

378. Jean-Yves Tissot and Clémentine Prieur. Bias correction for the estimation of sensitivity indices based on random balance designs. *Reliability Engineering & System Safety*, 107:205–213, 2012.

379. S Tarantola, D Gatelli, SS Kucherenko, W Mauntz, et al. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability Engineering & System Safety*, 92(7):957–960, 2007.

380. Alexandre Janon, Thierry Klein, Agnes Lagnoux, Maëlle Nodet, and Clémentine Prieur. Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: Probability and Statistics*, 18:342–364, 2014.

381. Sebastien Da Veiga, Francois Wahl, and Fabrice Gamboa. Local polynomial estimation for sensitivity analysis on models with correlated inputs. *Technometrics*, 51(4):452–463, 2009.

382. Bertrand Iooss, Alexandre Janon, and Gilles Pujol. Package 'sensitivity', version 1.17.1. *The Comprehensive R Archive Network*, pages 1–114, Feb 2020. URL https://cran.r-project.org/web/packages/sensitivity/.

383. Liang Liang, Fanwei Kong, Caitlin Martin, Thuy Pham, Qian Wang, James Duncan, and Wei Sun. Machine learning-based 3-D geometry reconstruction and modeling of aortic valve deformation using 3-D computed tomography images. *International Journal for Numerical Methods in Biomedical Engineering*, 33(5): e2827, 2017. doi: 10.1002/cnm.2827. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2827. e2827 CNM-Jan-16-0019.R1.

384. Nabil R. Adam, Robert Wieder, and Debopriya Ghosh. Data science, learning, and applications to biomedical and health sciences. *Annals of the New York Academy of Sciences*, 1387(1):5–11, 2017. doi: 10.1111/nyas.13309. URL https://nyaspubs.onlinelibrary.wiley.com/doi/abs/10.1111/nyas.13309.

385. N. Pozin, S. Montesantos, I. Katz, M. Pichelin, I. Vignon-Clementel, and C. Grandmont. Predicted airway obstruction distribution based on dynamical lung ventilation data: A coupled modeling-machine learning methodology. *International Journal for Numerical Methods in Biomedical Engineering*, 34(9):e3108, 2018. doi: 10.1002/cnm.3108. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3108. e3108 CNM-Jul-17-0180.R2.

386. Tobias Koeppl, Gabriele Santin, Bernard Haasdonk, and Rainer Helmig. Numerical modelling of a peripheral arterial stenosis using dimensionally reduced models and kernel methods. *International Journal for Numerical Methods in Biomedical Engineering*, 34(8):e3095, 2018. doi: 10.1002/cnm.

3095. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3095. e3095 cnm.3095.

387. Robert Koprowski and Kenneth R. Foster. Machine learning and medicine: book review and commentary. *BioMedical Engineering OnLine*, 17(1):17, Feb 2018. ISSN 1475-925X. doi: 10.1186/s12938-018-0449-9. URL https://doi.org/10.1186/s12938-018-0449-9.

388. Zixuan Cang and Guo-Wei Wei. Integration of element specific persistent homology and machine learning for protein-ligand binding affinity prediction. *International Journal for Numerical Methods in Biomedical Engineering*, 34(2):e2914, 2018. doi: 10.1002/cnm.2914. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2914. e2914 cnm.2914.

389. Yuanming Luo, Zhiwei Fan, Stephen Baek, and Jia Lu. Machine learning-aided exploration of relationship between strength and elastic properties in ascending thoracic aneurysm. *International Journal for Numerical Methods in Biomedical Engineering*, 34(6):e2977, 2018. doi: 10.1002/cnm.2977. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.2977. e2977 cnm.2977.

390. Myriam Cilla, Ignacio Pérez-Rey, Miguel Angel Martínez, Estefania Peña, and Javier Martínez. On the use of machine learning techniques for the mechanical characterization of soft biological tissues. *International Journal for Numerical Methods in Biomedical Engineering*, 34(10):e3121, 2018. doi: 10.1002/cnm.3121. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3121. e3121 cnm.3121.

391. Evgin Goceri. Diagnosis of Alzheimer's disease with Sobolev gradient-based optimization and 3D convolutional neural network. *International Journal for Numerical Methods in Biomedical Engineering*, 35(7):e3225, 2019. doi: 10.

1002/cnm.3225. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3225. e3225 CNM-Feb-19-0064.R1.

392. Bei Zhang, Ning Dai, Sukun Tian, Fulai Yuan, and Qing Yu. The extraction method of tooth preparation margin line based on s-octree cnn. *International Journal for Numerical Methods in Biomedical Engineering*, 35(10):e3241, 2019. doi: 10.1002/cnm.3241. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3241. e3241 CNM-Feb-19-0059.R1.

393. Jianhong Chen, Huang Huang, Wenrui Hao, and Jinchao Xu. A machine learning method correlating pulse pressure wave data with pregnancy. *International Journal for Numerical Methods in Biomedical Engineering*, 36(1):e3272, 2020. doi: 10.1002/cnm.3272. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3272.

394. Duc Duy Nguyen and Guo-Wei Wei. DG-GL: Differential geometry-based geometric learning of molecular datasets. *International Journal for Numerical Methods in Biomedical Engineering*, 35(3):e3179, 2019. doi: 10.1002/cnm.3179. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3179. e3179 cnm.3179.

395. Florian Joly, Gilles Soulez, Simon Lessard, Claude Kauffmann, and Irene Vignon-Clementel. A cohort longitudinal study identifies morphology and hemodynamics predictors of abdominal aortic aneurysm growth. *Annals of Biomedical Engineering*, 48(2):606–623, 2020.

396. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning. With applications in R.*, volume 103. New York, NY: Springer, 2013. ISBN 978-1-4614-7137-0/hbk; 978-1-4614-7138-7/ebook.

397. Yalin Başţanlar and Mustafa Özuysal. *Introduction to Machine Learning*, volume 1107 of *Methods in Molecular Biology*, chapter 7, pages 105–128. Humana Press,

Totowa, NJ, 2014. ISBN 978-1-62703-748-8. doi: 10.1007/978-1-62703-748-8_7. URL https://doi.org/10.1007/978-1-62703-748-8_7. (M.Yousef and J. Allmer, eds).

398. Joseph A. Cruz and David S. Wishart. Applications of machine learning in cancer prediction and prognosis. *Cancer Informatics*, 2:117693510600200030, 2006. doi: 10.1177/117693510600200030. URL https://doi.org/10.1177/117693510600200030.

399. Jeremy C. Weiss, S. Natarajan, Peggy L. Peissig, Catherine A. McCarty, and D. Page. Machine learning for personalized medicine: Predicting primary myocardial infarction from electronic health records. *AI Magazine*, 33(4), 12 2012. doi: https://doi.org/10.1609/aimag.v33i4.2438. URL https://aaai.org/ojs/index.php/aimagazine/article/view/2438.

400. Jeremy C. Weiss, F. Kuusisto, K. Boyd, J. Liu, and D. Page. Machine learning for treatment assignment: improving individualized risk attribution. *American Medical Informatics Association (AMIA) Annual Symposium*, 1306(15), 2015.

401. Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, and Dimitrios I. Fotiadis. Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13:8 – 17, 2015. ISSN 2001-0370. doi: https://doi.org/10.1016/j.csbj.2014.11.005. URL http://www.sciencedirect.com/science/article/pii/S2001037014000464.

402. Benjamin A. Goldstein, Ann Marie Navar, and Rickey E. Carter. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *European Heart Journal*, 38(23):1805–1814, 07 2016. ISSN 0195-668X. doi: 10.1093/eurheartj/ehw302. URL https://dx.doi.org/10.1093/eurheartj/ehw302.

403. Gabriel Pui Cheong Fung, J. X. Yu, Hongjun Lu, and P. S. Yu. Text classification without negative examples revisit. *IEEE Transactions on Knowledge and Data Engineering*, 18(1):6–20, Jan 2006. ISSN 1041-4347. doi: 10.1109/TKDE.2006.16.

404. H. Al-Mubaid and S. A. Umair. A new text categorization technique using distributional clustering and learning logic. *IEEE Transactions on Knowledge and Data Engineering*, 18(9):1156–1165, Sep. 2006. ISSN 1041-4347. doi: 10.1109/TKDE.2006.135.

405. Wenyuan Dai, Gui-Rong Xue, Qiang Yang, and Yong Yu. Co-clustering based classification for out-of-domain documents. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, pages 210–219, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-609-7. doi: 10.1145/1281192.1281218. URL `http://doi.acm.org/10.1145/1281192.1281218`.

406. K. Sarinnapakorn and M. Kubat. Combining subclassifiers in text categorization: A dst-based solution and a case study. *IEEE Transactions on Knowledge and Data Engineering*, 19(12):1638–1651, Dec 2007. ISSN 1041-4347. doi: 10.1109/TKDE.2007.190663.

407. Rajat Raina, Andrew Y. Ng, and Daphne Koller. Constructing Informative Priors Using Transfer Learning. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, pages 713–720, New York, NY, USA, 2006. ACM. ISBN 1-59593-383-2. doi: 10.1145/1143844.1143934. URL `http://doi.acm.org/10.1145/1143844.1143934`.

408. John Blitzer, Ryan McDonald, and Fernando Pereira. Domain adaptation with structural correspondence learning. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, EMNLP '06, pages 120–128, Stroudsburg, PA, USA, 2006. Association for Computational Linguistics. ISBN

1-932432-73-6. URL http://dl.acm.org/citation.cfm?id=1610075.1610094.

409. Sinno Pan, Vincent Wenchen Zheng, Qiang Yang, and Derek Hao Hu. Transfer learning for wifi-based indoor localization. *AAAI Workshop - Technical Report*, 01 2008.

410. Matthew E. Taylor and Peter Stone. Cross-domain transfer for reinforcement learning. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pages 879–886, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-793-3. doi: 10.1145/1273496.1273607. URL http://doi.acm.org/10.1145/1273496.1273607.

411. Jan Ramon, Kurt Driessens, and Tom Croonenborghs. Transfer learning in reinforcement learning problems through partial policy recycling. In *Proceedings of the 18th European Conference on Machine Learning*, ECML '07, pages 699–707, Berlin, Heidelberg, 2007. Springer-Verlag. ISBN 978-3-540-74957-8. doi: 10.1007/978-3-540-74958-5_70. URL http://dx.doi.org/10.1007/978-3-540-74958-5_70.

412. Matthew E. Taylor and Peter Stone. Transfer learning for reinforcement learning domains: A survey. *J. Mach. Learn. Res.*, 10:1633–1685, December 2009. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=1577069.1755839.

413. Alessandro Lazaric. *Transfer in Reinforcement Learning: A Framework and a Survey*, chapter 5, pages 143–173. Springer Berlin Heidelberg, Berlin, Heidelberg, first edition, 2012. ISBN 978-3-642-27645-3. doi: 10.1007/978-3-642-27645-3_5. URL https://doi.org/10.1007/978-3-642-27645-3_5. (M. Wiering and M. van Otterlo, eds).

414. Yusen Zhan and Matthew E. Taylor. Online transfer learning in reinforcement learning domains. *CoRR*, abs/1507.00436, 2015. URL http://arxiv.org/abs/1507.00436.

415. Vincent Dumoulin and Francesco Visin. A guide to convolution arithmetic for deep learning. *arXiv e-prints*, art. arXiv:1603.07285, March 2016.

416. Nigel Williams, Sebastian Zander, and Grenville Armitage. A Preliminary Performance Comparison of Five Machine Learning Algorithms for Practical IP Traffic Flow Classification. *SIGCOMM Comput. Commun. Rev.*, 36(5):5–16, October 2006. ISSN 0146-4833. doi: 10.1145/1163593.1163596. URL http://doi.acm.org/10.1145/1163593.1163596.

417. Xindong Wu, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, Angus Ng, Bing Liu, Philip S. Yu, Zhi-Hua Zhou, Michael Steinbach, David J. Hand, and Dan Steinberg. Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1):1–37, Jan 2008. ISSN 0219-3116. doi: 10.1007/s10115-007-0114-2. URL https://doi.org/10.1007/s10115-007-0114-2.

418. Chih-Fong Tsai, Yu-Feng Hsu, Chia-Ying Lin, and Wei-Yang Lin. Intrusion detection by machine learning: A review. *Expert Systems with Applications*, 36 (10):11994 – 12000, 2009. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2009.05.029. URL http://www.sciencedirect.com/science/article/pii/S0957417409004801.

419. John Joseph Valletta, Colin Torney, Michael Kings, Alex Thornton, and Joah Madden. Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 124:203 – 220, 2017. ISSN 0003-3472. doi: https://doi.org/10.1016/j.anbehav.2016.12.005. URL http://www.sciencedirect.com/science/article/pii/S0003347216303360.

420. N. R. Draper and H. Smith. *Applied Regression Analysis*. Wiley, London, 2014.

421. Trevor Hastie, Robert Tibshirani, and Jerome Firedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer-Verlag, New York, 2017. ISBN 9780387952840. URL https://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12.pdf.

422. G Mclachlan, Kim-Anh Do, and Christophe Ambroise. *Discriminant Analysis*, chapter 6, pages 185–220. John Wiley & Sons, Ltd, 2005. ISBN 9780471728429. doi: 10.1002/047172842X.ch6. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/047172842X.ch6.

423. Max Kuhn and Kjell Johnson. *Applied Predictive Modeling*. Springer, New York, 2013. ISBN 978-1-4614-6848-6. doi: 10.1007/978-1-4614-6849-3.

424. Leo Breiman and Philip Spector. Submodel selection and evaluation in regression. the x-random case. *International Statistical Review / Revue Internationale de Statistique*, 60(3):291–319, 1992. ISSN 03067734, 17515823. URL http://www.jstor.org/stable/1403680.

425. Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, pages 1137–1143, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc. ISBN 1-55860-363-8. URL http://dl.acm.org/citation.cfm?id=1643031.1643047.

426. Hao Liu, Fuyou Liang, Jasmin Wong, Takashi Fujiwara, Wenjing Ye, Ken-iti Tsubota, and Michiko Sugawara. Multi-scale modeling of hemodynamics in the cardiovascular system. *Acta Mechanica Sinica*, 31(4):446–464, 2015.

427. J Alastruey, KH Parker, J Peiró, and SJ Sherwin. Lumped parameter outflow models for 1-d blood flow simulations: effect on pulse waves and parameter estimation. *Communications in Computational Physics*, 4(2):317–336, 2008.

428. Liesbeth Taelman, Joris Degroote, Pascal Verdonck, Jan Vierendeels, and Patrick Segers. Modeling hemodynamics in vascular networks using a geometrical multiscale approach: numerical aspects. *Annals of biomedical engineering*, 41(7): 1445–1458, 2013.

429. WW Chen, H Gao, XY Luo, and NA Hill. Study of cardiovascular function using a coupled left ventricle and systemic circulation model. *Journal of biomechanics*, 49(12):2445–2454, 2016.

430. Yanhang Zhang, Victor H Barocas, Scott A Berceli, Colleen E Clancy, David M Eckmann, Marc Garbey, Ghassan S Kassab, Donna R Lochner, Andrew D McCulloch, Roger Tran-Son-Tay, et al. Multi-scale modeling of the cardiovascular system: disease development, progression, and clinical intervention. *Annals of biomedical engineering*, 44(9):2642–2660, 2016.

431. Alfio Quarteroni, Alessandro Veneziani, and Christian Vergara. Geometric multiscale modeling of the cardiovascular system, between theory and practice. *Computer Methods in Applied Mechanics and Engineering*, 302:193–252, 2016.

432. Leopold Grinberg, Elizabeth Cheever, Tomer Anor, Joseph R Madsen, and GE Karniadakis. Modeling blood flow circulation in intracranial arterial networks: a comparative 3D/1D simulation study. *Annals of biomedical engineering*, 39(1): 297–309, 2011.

433. Luca Formaggia, Fabio Nobile, Alfio Quarteroni, and Alessandro Veneziani. Multiscale modelling of the circulatory system: a preliminary analysis. *Computing and visualization in science*, 2(2-3):75–83, 1999. URL `https://doi.org/10.1007/s007910050030`.

434. J Arimon Alastruey. *Numerical modelling of pulse wave propagation in the cardiovascular system: development, validation and clinical applications.* PhD thesis, University of London, 2006.

435. Kristen DeVault, Pierre A Gremaud, Vera Novak, Mette S Olufsen, Guillaume Vernieres, and Peng Zhao. Blood flow in the circle of willis: modeling and calibration. *Multiscale Modeling & Simulation*, 7(2):888–909, 2008.

436. Pablo J Blancoa and Raúl A Feijóoa. A 3D-1D-0D Computational model for the entire cardiovascular system. *Mecánica Computacional*, 29(59):5887–5911, 2010.

437. PJ Blanco, RA Feijóo, and SA Urquiza. A unified variational approach for coupling 3D–1D models and its blood flow applications. *Computer Methods in Applied Mechanics and Engineering*, 196(41):4391–4410, 2007.

438. S Moore, T David, JG Chase, J Arnold, and J Fink. 3D models of blood flow in the cerebral vasculature. *Journal of biomechanics*, 39(8):1454–1463, 2006.

439. Sandeep Gutta, Qi Cheng, and Bruce A Benjamin. Control mechanism modeling of human cardiovascular-respiratory system. In *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 918–922. IEEE, 2015.

440. FY Liang. *An integrated computational study of multi-scale hemodynamics and multi-mechanism physiology in human cardiovascular system.* PhD thesis, PhD thesis, Chiba University, Artificial System Science Department, 2007.

441. Mauro Ursino. Interaction between carotid baroregulation and the pulsating heart: a mathematical model. *American Journal of Physiology-Heart and Circulatory Physiology*, 275(5):H1733–H1747, 1998.

442. Jing Bai, Hongli Lu, Jupeng Zhang, Bing Zhao, and Xiaoqiang Zhou. Optimization and mechanism of step–leap respiration exercise in treating of cor pulmonale. *Computers in biology and medicine*, 28(3):289–307, 1998.

443. Bram W Smith, Steen Andreassen, Geoffrey M Shaw, Per L Jensen, Stephen E Rees, and J Geoffrey Chase. Simulation of cardiovascular system diseases by including the autonomic nervous system into a minimal model. *Computer methods and programs in biomedicine*, 86(2):153–160, 2007.

444. Antonio Albanese, Limei Cheng, Mauro Ursino, and Nicolas W Chbat. An integrated mathematical model of the human cardiopulmonary system: model development. *American Journal of Physiology-Heart and Circulatory Physiology*, 310(7):H899–H921, 2016.

445. Paulo Roberto Trenhago, Luciano Gonçalves Fernandes, Lucas Omar Müller, Pablo Javier Blanco, and Raúl Antonino Feijóo. An integrated mathematical model of the cardiovascular and respiratory systems. *International journal for numerical methods in biomedical engineering*, 32(1):e02736, 2016.

446. Fabrizia Salerni, Rodolfo Repetto, Alon Harris, Peter Pinsky, Christophe Prud'homme, Marcela Szopos, and Giovanna Guidoboni. Biofluid modeling of the coupled eye-brain system and insights into simulated microgravity conditions. *PloS one*, 14(8), 2019.

447. Jan Poleszczuk, Malgorzata Debowska, Wojciech Dabrowski, Alicja Wojcik-Zaluska, Wojciech Zaluska, and Jacek Waniewski. Patient-specific pulse wave propagation model identifies cardiovascular risk characteristics in hemodialysis patients. *PLoS computational biology*, 14(9):e1006417, 2018.

448. Zheng Duanmu, Min Yin, Xueling Fan, Xilan Yang, and Xiaoyu Luo. A patient-specific lumped-parameter model of coronary circulation. *Scientific reports*, 8(1): 1–10, 2018.