



UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI INGEGNERIA INDUSTRIALE

**TESI DI LAUREA MAGISTRALE IN
INGEGNERIA CHIMICA E DEI PROCESSI INDUSTRIALI**

**UN APPROCCIO INGEGNERISTICO
ALL'IDENTIFICAZIONE DI MODELLI FISIOLGICI
PER LA DESCRIZIONE DELLA MALATTIA
DI VON WILLEBRAND**

Relatore: Prof. Fabrizio Bezzo

Correlatore: Dr. Federico Galvanin

Laureando: ANDREA MONTE

ANNO ACCADEMICO 2012-2013

*“Remember that all models are wrong;
the practical question is how wrong do
they have to be to not be useful.”*

George E. P. Box

Riassunto

La malattia di von Willebrand è una patologia ereditaria del sistema di coagulazione sanguigno caratterizzata da un elevato numero di varianti. La diagnosi è resa ancor più complessa dalla grande variabilità intra-individuale. La disponibilità di modelli fisiologici affidabili potrebbe essere di aiuto per sviluppare una diagnostica più rapida ed efficace.

In questo lavoro alcuni modelli proposti sono presi in considerazione per proporre delle tecniche che ne consentono l'identificabilità parametrica. Inoltre l'analisi di sensitività dei modelli indica come sia possibile progettare dei test clinici altamente informativi con campionamenti meno onerosi per il paziente. È proposta infine una procedura di diagnosi mediante modello che, attraverso delle stime parametriche individuali basate su dati di elettroforesi, permette di identificare la tipologia sofferta dal paziente.

3.1	Analisi dei dati disponibili	33
3.2	Formulazione della struttura dei modelli candidati	36
3.3	Modelli proposti	37
3.4	Stime preliminari dei modelli 6M e 7	40
3.4.1	Stime effettuate per i soggetti normali 0 e non-0	41
3.4.2	Stime effettuate per i pazienti affetti da VWD	44
3.5	Analisi di sensitività	48
3.5.1	Varianti del modello 6M	52
3.6	Analisi di informazione e correlazione	54
3.7	Osservazioni sui modelli	58
CAPITOLO 4 Diagnosi della malattia di von Willebrand mediante modello		59
4.1	Diagnosi mediante il modello 6M	59
4.2	Dati da analisi di gel-elettroforesi	61
4.3	Utilizzo dei dati disponibili	64
4.4	Risultati delle stime per la diagnosi della VWD	65
4.4.1	Diagnosi da dati VWF:AG e CB individuali	72
4.5	Stima dei parametri individuali	74
4.6	Analisi di sensitività e di informazione sulle frazioni	78
4.7	Riparametrizzazione	82
4.8	Simulazione stocastica	85
4.9	Conclusioni	88
CONCLUSIONI		89
NOMENCLATURA		91
APPENDICE A Stime parametriche ottenute per la diagnosi della VWD		95
APPENDICE B Stime parametriche ottenute con il set riparametrizzato		101
RINGRAZIAMENTI		111

Introduzione

La malattia di von Willebrand (VWD) è considerata la più comune patologia ereditaria del sangue che interessa il processo coagulativo. È una disfunzione genetica caratterizzata da difetti qualitativi, quantitativi o dalla totale assenza nel plasma sanguigno del fattore di von Willebrand (VWF), una glicoproteina multimerica di grandi dimensioni, che svolge diversi ruoli nel processo di coagulazione sanguigno. La malattia è molto eterogenea e presenta una grande varietà di forme i cui sintomi possono variare da sporadici sanguinamenti spontanei da piccole lesioni cutanee o delle mucose, a severi e prolungati episodi emorragici simili a casi di emofilia, fino in casi più rari a problemi di tipo trombotico.

Semplici modelli farmacocinetici sono stati utilizzati in studi recenti (Casonato *et al.*, 2010), per investigare sia le cinetiche di distribuzione che lo smaltimento del VWF nel plasma sanguigno per distinte classi di pazienti, fornendo indicazioni essenziali per lo studio e la descrizione della malattia. Tuttavia non è ancora stato sviluppato un modello matematico affidabile, basato sulla fisiologia del singolo soggetto, che descriva in modo quantitativo la distribuzione dei multimeri di VWF nel tempo. La possibilità di avere a disposizione un modello matematico rappresenta un passo in avanti per: aumentare la comprensione dei complessi fenomeni fisiologici che avvengono nei pazienti affetti da VWD; fornire gli strumenti necessari in modo tale da proporre un protocollo per una diagnosi più sicura e veloce; consentire delle simulazioni in silico per aprire la strada verso una terapia più efficiente.

L'obiettivo principale di questo lavoro di tesi è quello di identificare, tramite un approccio *Process Systems Engineering* (PSE), un modello meccanicistico per lo studio della VWD. Le tecniche PSE hanno da tempo dimostrato di essere altamente efficaci nello sviluppo di modelli e identificazione di molti sistemi complessi nell'industria di processo, più recentemente è stato dimostrato inoltre che possono essere applicate con successo a diversi sistemi biologici e fisiologici.

La Tesi è strutturata come segue.

Nel Capitolo 1, dopo una breve introduzione sul complesso processo di coagulazione sanguigna, viene descritto il VWF e le diverse disfunzioni che caratterizzano la VWD nelle varie tipologie;

successivamente si illustrano i diversi metodi diagnostici utilizzati in campo medico, ponendo l'attenzione sulle problematiche legate alla difficoltà di una precisa diagnosi.

Il Capitolo 2 contiene la base teorica degli strumenti matematici e statistici utilizzati per le attività di modellazione, validazione ed identificabilità parametrica di modelli fisiologici. Alla fine del Capitolo si presentano i fondamenti della progettazione degli esperimenti basata sul modello per contestualizzare l'attività di modellazione.

Nel Capitolo 3 si discutono i modelli proposti per la descrizione della malattia di von Willebrand, descrivendo le problematiche strutturali e di identificabilità parametrica legate alla descrizione dei meccanismi delle diverse tipologie della malattia.

Il Capitolo 4 è incentrato sull'utilizzo del modello matematico per la diagnosi della malattia di von Willebrand, tramite delle stime parametriche in riferimento ai diversi dati sperimentali disponibili per ciascun paziente. Infine, dopo una discussione dei risultati ottenuti, si presenta una riparametrizzazione del modello per risolvere i problemi di identificabilità in funzione di possibili sviluppi futuri tramite tecniche MBD_{oE}.

La tesi è conclusa da una serie di considerazioni riassuntive sul lavoro svolto e sui risultati ottenuti. Sono inoltre discusse alcune questioni critiche da affrontare in una futura ricerca sull'argomento.

Capitolo 1

La malattia di von Willebrand

Il fattore di von Willebrand (VWF) gioca un ruolo molto importante all'interno del processo di emostasi. Per comprendere al meglio cosa sia tale fattore, quali siano le sue funzioni e discernere agevolmente le diverse forme della malattia di von Willebrand (VWD), sarà introdotto brevemente l'intero processo di emostasi; in seguito si descriverà la struttura e il processo di biosintesi del VWF, infine la patologia, le sue diverse tipologie ed i metodi diagnostici.

1.1 Introduzione

Nel 1924 Erik von Willebrand, un medico finlandese, iniziò a studiare una patologia ereditaria che affliggeva le famiglie consanguinee dell'isola di Åland, situata nello stretto Golfo di Bothnia che separa la Svezia dalla Finlandia. Il primo soggetto studiato fu una bambina di cinque anni con seri problemi emorragici sin dalla nascita. Tre delle sue sorelle morirono prima di raggiungere quattro anni, mentre i genitori, due fratelli e i restanti parenti di entrambi i sessi presentavano problemi meno gravi, i quali si acuiscono qualora si presentavano casi di omozigosi. von Willebrand distinse le condizioni tra casi di anemia ed altri disordini congeniti allora conosciuti, come l'emofilia A, ma non fu in grado di determinare se la causa delle forti emorragie fosse dovuta a difetti nelle piastrine, del plasma sanguigno e nei vasi sanguigni.

Oggi la VWD è considerata la più comune patologia ereditaria del sangue che interessa il processo coagulativo. È una disfunzione genetica autosomica dominante caratterizzata da difetti qualitativi, quantitativi o dalla totale assenza di VWF, una glicoproteina di grandi dimensioni che svolge diversi ruoli nel processo di coagulazione sanguigno. Meno frequentemente la malattia può essere acquisita (sindrome di VW acquisita). I sintomi clinici più comuni includono: sanguinamenti di naso, ematomi cutanei, emorragie derivanti da piccole lesioni cutanee, nelle mucose o nel tratto intestinale, menorragia ed eccessiva perdita di sangue in seguito traumi, interventi chirurgici e parti (Lillicrap, 2007).

Studi condotti sia in Italia (Rodeghiero *et al.*, 1987) che negli U.S.A. (Werner *et al.*, 1993) stimano la frequenza di comparsa della VWD attorno all'1% a livello mondiale. Dati clinici più recenti provenienti da ospedali terzi (Goodeve, 2010) suggeriscono che, basandosi su una

popolazione di 5.8 miliardi, esistano circa 580000 soggetti nel mondo che presentano i sintomi della VWD, cioè 100 casi per milione, ed approssimativamente l'80% di queste persone vive in paesi sviluppati o in via di sviluppo.

Negli ultimi anni sono stati fatti passi considerevoli nella caratterizzazione delle specifiche molecole responsabili dei difetti del VWF, tuttavia il complesso meccanismo che guida la distribuzione, la rimozione e la sopravvivenza del VWF nel plasma non è ancora chiaro agli occhi della comunità scientifica. Risulta evidente che la formulazione di un modello, che descriva l'insieme dei meccanismi sopraelencati, e lo studio tramite un approccio Process System Engineering (PSE), possa essere di notevole aiuto per la comprensione della malattia in se, ma anche come supporto medico in fase di diagnosi e cura o semplicemente per effettuare test e simulazioni in silico prima di coinvolgere i pazienti in gravosi test in vivo.

1.2 L'emostasi

Con il termine emostasi si intendono una serie di processi biochimici e cellulari, che hanno lo scopo di mantenere la fluidità del sangue all'interno dei vasi sanguigni e di indurre la formazione del coagulo emostatico a seguito di un danno o trauma vascolare. È un meccanismo di difesa caratterizzato da tre fasi (Hoffman e Monroe, 2001):

- **fase vascolare:** vasocostrizione del vaso interessato dal trauma;
- **emostasi primaria:** adesione, attivazione e aggregazione piastrinica;
- **emostasi secondaria:** produzione dei fattori di coagulazione e attivazione della cascata coagulativa con conseguente formazione di un coagulo.

Una volta terminati i processi riparativi, si instaura il processo di retroazione del coagulo e fibrinolisi, un insieme di meccanismi di contro-regolazione che bilanciano il sistema di coagulazione limitando la formazione del tappo emostatico alla sede della lesione.

Un danno vascolare induce vasocostrizione, cioè contrazione dovuta a meccanismi neurogeni riflessi delle cellule della tonaca media, la parte più spessa del vaso che gli conferisce elasticità e contrattilità essendo costituita da fibrocellule muscolari lisce, e potenziata dalla secrezione locale da parte dell'endotelio di endotelina, un polipeptide ad azione vasocostrittrice che favorisce la proliferazione di diversi tipi di cellule della parete vascolare, e di serotonina da parte delle piastrine. L'effetto è transitorio e il sanguinamento continuerebbe se non venisse attivato il sistema delle piastrine e la cascata coagulativa. Le cellule endoteliali a causa della lesione secernono il fattore di von Willebrand (VWF), una proteina che si dispone presso la lesione e

permette l'adesione piastrinica mediando l'interazione tra le piastrine e la matrice extracellulare (ECM) esposta che è trombogena. Le piastrine vi si legano tramite la glicoproteina Ib (GpIb) e a sua volta il VWF si associa al collagene della ECM.

Entro qualche minuto le piastrine iniziano ad aderire al VWF e cambiano forma, da discoidale a piatta, aumentando la loro superficie; esse rilasciano poi i loro granuli secretori contenenti prevalentemente adenosindifosfato (ADP) e trombossano (TXA₂). Queste due sostanze fungono da chemochine¹ per altre piastrine che si accumulano presso la lesione apponendosi sulle altre già presenti formando il tappo emostatico primario (emostasi primaria).

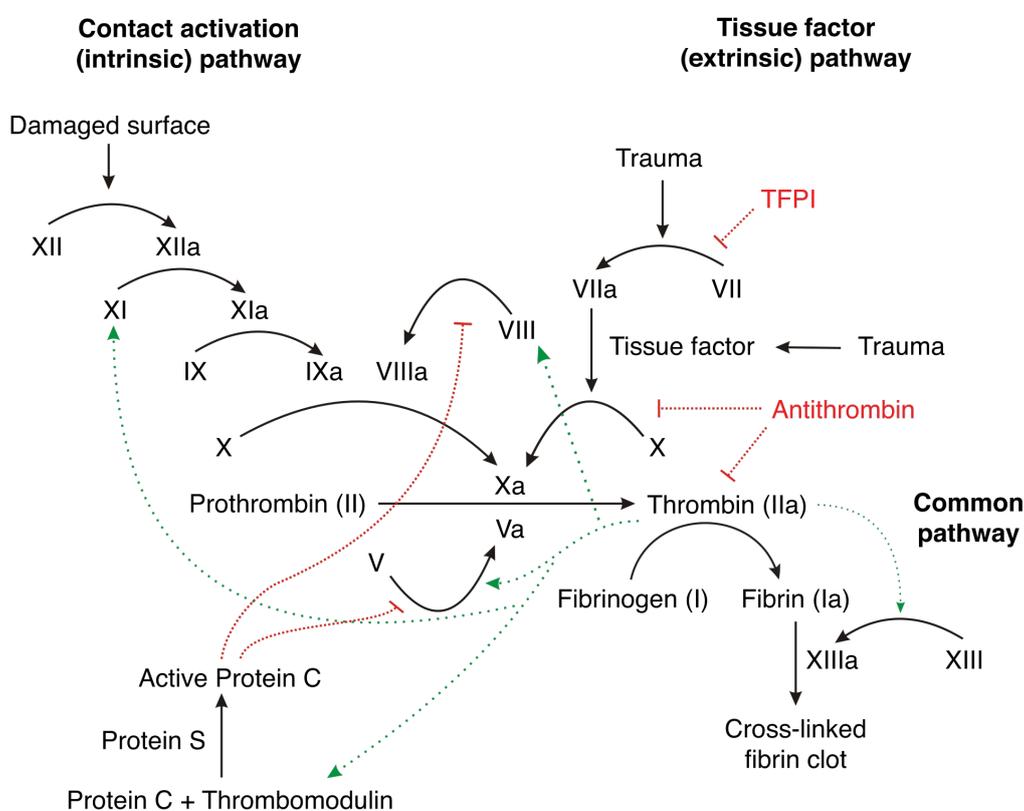


Figura 1.1: Processi della cascata coagulativa. Si noti come la trombina attivata instauri un secondo loop di attivazione rappresentato dalle linee tratteggiate in verde. Le linee rosse invece descrivono i meccanismi di controregolazione della cascata coagulativa tramite gli anticoagulanti proteine C ed S, l'inibitore del fattore tissutale (TFPI) e l'antitrombina (Hoffman e Monroe, 2001).

Successivamente si avrà l'emostasi secondaria, fase caratterizzata dalla stabilizzazione dell'aggregato grazie all'attivazione della fibrina. In questo frangente l'endotelio secerne il fattore tissutale (TF) o tromboplastina tissutale, altrimenti indicato come fattore III, una proteina pre-

¹Le chemochine sono un gruppo di proteine a basso peso molecolare della famiglia delle citochine. La loro funzione principale consiste nell'attivazione e nel reclutamento (chemiotassi) dei leucociti per regolarne il traffico nell'organismo.

sente nel tessuto subendoteliale, nelle piastrine, e nei leucociti necessaria per la formazione della trombina dal suo precursore: la protrombina. Le piastrine successivamente espongono sulla loro superficie particolari fosfolipidi i quali favoriscono l'adesione del TF. Il processo di attivazione della fibrina segue invece un meccanismo a cascata, descritto in Figura 1.1, che vede la partecipazione di molti più fattori coagulanti, che in totale sono 13, alcuni dei quali sono degli enzimi, altri invece delle glicoproteine pro-enzimatiche di piccole dimensioni.

La fibrina si trova normalmente sotto forma di fibrinogeno, glicoproteina circolante e solubile del plasma sanguigno, che non può dar luogo ad un aggregato. Per far sì che il fibrinogeno venga attivato esistono due vie, una intrinseca ed una estrinseca. Queste due vie differiscono tra di loro principalmente per l'agente iniziale che le attiva e il numero di fattori coinvolti nella cascata, ma elementi dell'una possono influenzare l'attivazione dell'altra. Le due vie si congiungono poi, originando la via comune, che ha inizio con l'attivazione del fattore X.

La via estrinseca, più rapida per il minor numero di fattori che vi prendono parte, viene attivata quando una lesione di un vaso sanguigno produce la liberazione, dalle cellule danneggiate, di fosfolipidi e del complesso proteico TF. Il fattore VII attivato si lega a sua volta al TF, il complesso così formato attiva in cascata il fattore X. La via intrinseca più lenta, si verifica quando il sangue entra a contatto con la ECM, in particolare con le macromolecole di collagene, ed innesca l'attivazione del fattore XII, o fattore di Hageman. Per taglio enzimatico il fattore XII attiva il fattore XI, il quale in cascata attiva il fattore IX nello stesso modo. Il fattore IX attivato si unisce al fattore VIII, il quale trasportato dal VWF, si scinde attivandosi (VIIIa) e provoca a sua volta l'attivazione del fattore X, ancora una volta per proteolisi. La via comune comincia con la formazione di un complesso tra fattore X attivato (Xa) e V attivato (Va), il quale attiva la protrombina (II), che similmente attiva il fibrinogeno (I). A questo punto il fibrinogeno si trova in uno stato solubile, quindi l'ulteriore intervento del fattore XIII attivato (XIIIa) permette la precipitazione della fibrina e la formazione del tappo piastrinico permanente.

In seguito vengono attivati dei meccanismi di contro-regolazione e rimozione che agiscono in due modi: i primi atti a controbilanciare il sistema della cascata coagulativa, tramite degli inibitori anticoagulanti, le proteine C ed S, l'atitrombina, il fattore tissulare di inibizione (TP-FI) e l'eparina, che bloccano il processo limitando la formazione del tappo emostatico, le vie in rosso di Figura 1.1; il secondo invece è la fibrinolisi, il quale degrada i complessi solubili di fibrina e rimuove il trombo al termine dei processi riparativi tramite il sistema enzimatico plasminogeno-plasmina.

1.3 Fisiologia del fattore di von Willebrand

Come visto nel paragrafo precedente, il VWF ha due funzioni essenziali (Sadler, 2005): media la fase iniziale dell'adesione delle piastrine al subendotelio dei vasi sanguigni che hanno subito lesioni e l'adesione tra piastrina e piastrina, in secondo luogo funge da vettore per il fattore VIII proteggendolo dalla degradazione.

Il VWF è una macromolecola glicoproteica presente nel plasma sanguigno sintetizzata dalle cellule endoteliali e dai megacariociti, precursori delle piastrine localizzati a livello del midollo osseo. Esso circola nel plasma sottoforma di un macropolimero multimerico ad elevato peso molecolare (milioni di dalton) costituito da singole unità fondamentali tenute insieme da numerosi legami disolfuro, dovuti dall'elevato contenuto di cisteina, (Figura 1.2).

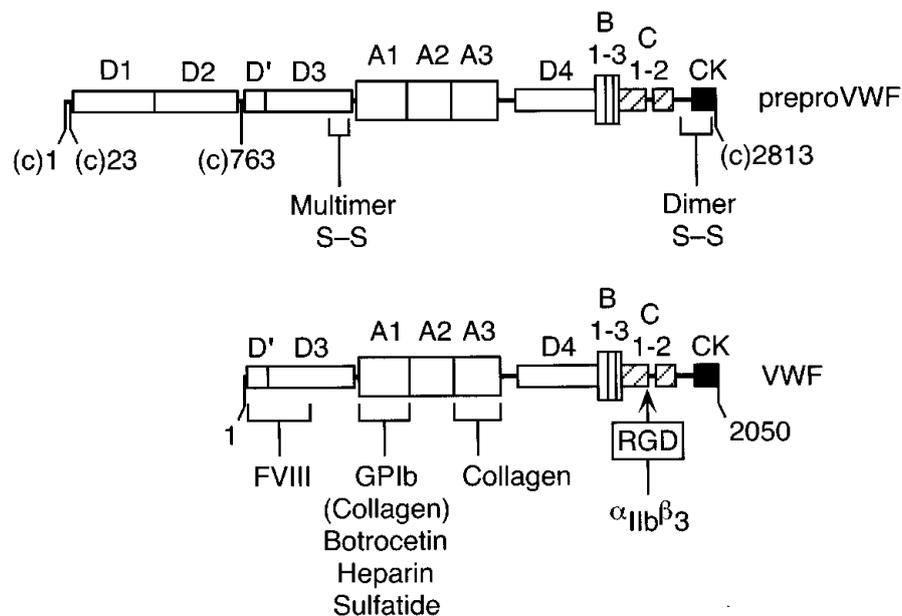


Figura 1.2: Struttura del prepro-VWF e del monomero VWF. Ci sono tre domini A, tre domini B, due domini C e quattro domini D. I domini A corrispondono alla regione del VWF a ridotto contenuto di cisteina, i domini B sono piccoli e contengono da 25 a 35 aminoacidi, mentre i domini C ne contengono da 116 a 119 e per finire i domini D ne contengono da 351 a 376. Molto importanti sono il dominio A3, responsabile del legame con il collagene di tipo I e III dei vasi danneggiati e il dominio A1, che causa l'aggregazione delle piastrine legandosi a specifici recettori di membrana (Sadler, 2005).

Il gene del VWF è localizzato nella banda 21 del cromosoma 12 e porta alla sintesi di una molecola che inizialmente presenta 2813 aminoacidi, detta prepro-fattore di von Willebrand, con un numero di sequenze ripetitive raggruppate in 4 diversi domini e distinte in tre regioni: la prima è un corto peptide segnale composto da 22 aminoacidi, la seconda una regione di 748 aminoacidi detta antigene II del VWF, la terza è la forma monomerica del VWF, Figura 1.2.

Nel reticolo endoplasmatico, le molecole di pre-pro-VWF formano dei dimeri e, sia il peptide segnale che l'antigene II del VWF vengono rimossi per proteolisi originando la subunità matura del VWF di 2050 aminoacidi. Il dimerico del VWF viene quindi trasferito nell'apparato del Golgi dove, mediante reazioni di O-glicosilazione e solfatazione, viene a formarsi la catena multimerica del VWF (Figura 1.3).

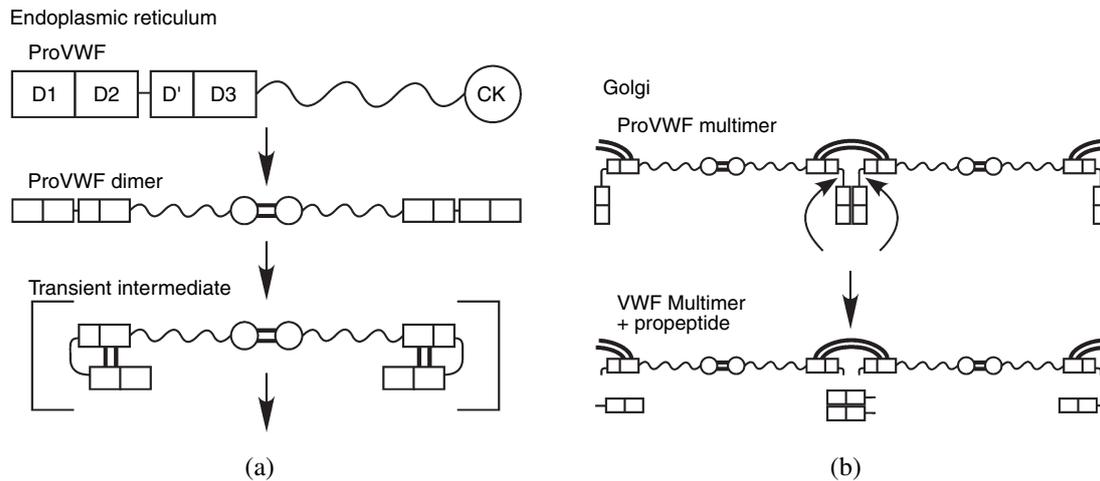


Figura 1.3: Processi di dimerizzazione del monomero VWF nel reticolo endoplasmatico (a) e di polimerizzazione nell'apparato del Golgi (b). Le doppie linee scure rappresentano i ponti disolfuro dovuti dall'amminoacido cisteina (Sadler, 2005).

Le molecole multimeriche definitive del VWF sono rilasciate dalle cellule endoteliali, processo detto di rilascio basale, oppure vengono depositate negli α -granuli delle piastrine o nei corpi di Weibel-Palade² pronte per l'utilizzo in seguito ai diversi stimoli coagulativi.

La funzione emostatica del VWF dipende fortemente dai legami che esso crea con le glicoproteine (GPIb) presenti sulla superficie delle piastrine, con il fattore VIII e con i siti attivi del tessuto connettivo esposto dal danno. Questa attività è modulata dall'organizzazione di diverse unità multimeriche derivanti dalla proteolisi ed eliminazione del VWF da parte di uno specifico enzima ADAMTS-13 metalloproteasi.

Una volta rilasciato nel flusso sanguigno il VWF ha un tempo di vita che si aggira tra le 12 e le 20 h, in questo tempo l'enzima ADAMTS-13 taglia nel dominio A2 i multimeri di VWF, attivati dal sufficiente shear-stress, creando una distribuzione di sub-unità a diversi pesi molecolari. Le forme ad alto peso molecolare sono considerate molto più attive ed efficienti rispetto ai multimeri di dimensioni minori, poiché hanno più affinità liganti ai siti attivi.

²I corpi di Weibel-Palade sono delle vescicole presenti nelle cellule endoteliali, e rappresentano il modo in cui viene immagazzinata la maggior parte di VWF.

La concentrazione stazionaria nel plasma di VWF dipende quindi dal bilancio tra sintesi ed eliminazione, mentre la distribuzione delle unità multimeriche deriva dai meccanismi competitivi di clearance e proteolisi da parte del ADAMTS-13.

1.4 Tipologie della VWD

In accordo con la classificazione della VWD si possono distinguere diverse tipologie di VWD (Sadler, 2002): tipo 1 e 3 sono difetti quantitativi di VWF, mentre il tipo 2 rappresenta difetti nell'assemblaggio del multimer VWF o una forte degradazione nella fase proteolitica da parte del ADAMTS-13. Inoltre la totale mancanza della fase proteolitica da parte del ADAMTS-13, e di conseguenza una maggior concentrazione di grossi multimeri di VWF, può causare la porpora trombotica trombocitopenica (TTP o sindrome di Moscovowitz).

1.4.1 VWD di tipo 1

La distribuzione dei multimeri VWF è normale in questa tipologia, ma vi è un deficit quantitativo di concentrazione plasmatica. I bassi livelli nel flusso sanguigno di VWF sono difficilmente rilevabili con precisione poiché la normale distribuzione di concentrazione del VWF presenta una variabilità molto ampia tra i soggetti sani. La concentrazione di VWF antigene si è osservato dipendere dalla tipologia di gruppo sanguigno AB0, dove i soggetti sani 0 presentano circa il 25% in meno rispetto ai pazienti non-0.

La VWD di tipo 1 è indipendente dall'attività enzimatica del ADAMTS-13, e l'eziologia è dovuta principalmente a mutazioni che riducono i livelli di VWF a causa di una ridotta sintesi³ e rilascio da parte delle cellule endoteliali e di una maggiore clearance plasmatica che riduce l'attività coagulante legata alle piastrine-VWF e al fattore VIII.

Studi condotti con la desmopressina hanno trovato tempi di vita del VWF intorno alle 1.3 h, quindi l'eliminazione è molto veloce.

Un caso estremo di questa variante è la VWD Vicenza, solitamente classificata sotto la tipologia 1, ma che può essere inclusa anche nel tipo 2M (vedi in seguito). I pazienti affetti da VWD Vicenza tipicamente presentano delle normali piastrine-VWF, normali concentrazioni di VWF, ma una distribuzione indistinguibile delle diverse unità multimeriche derivanti dalla proteolisi, producendo di conseguenza una rapida clearance e rapidi consumi del VWF. La VWD Vicenza è dovuta ad una mutazione nel monomero VWF del dominio D3 (Figura 1.2).

³A causa di una riduzione di trasporto del pro-VWF e quindi una minore dimerizzazione e successiva polimerizzazione nel Golgi.

1.4.2 VWD di tipo 2

In questa variante solo i multimeri VWF di grandi dimensioni hanno attività emostatica e la loro assenza selettiva causa problemi emorragici. Questo difetto qualitativo accomuna le varianti della VWD di tipo 2. La distribuzione dei multimeri di VWF nel plasma dipende dalla competizione delle fasi di clearance e proteolisi da parte del ADAMTS-13, ma può essere anche causata da mutazioni nel VWF, nelle GPIb delle piastrine e anche del ADAMTS-13.

VWD 2A

L'assenza di multimeri VWF di maggior peso molecolare nella VWD 2A è causata da una proteolisi prematura dei polimeri o dall'incapacità di assemblaggio degli stessi. Ciò porta ad aver concentrazioni più elevate sia nel plasma che nelle piastrine di multimeri a basso peso molecolare questo porterà ad un fallimento della fase di aggregazione piastrinica, causando problemi coagulativi.

VWD 2B

La VWD di tipo 2B è caratterizzata da una maggiore affinità delle piastrine a legarsi con le forme anomale di VWF ad elevato peso molecolare, e le restanti piastrine tenderanno a formare aggregati intravascolari i quali, a seguito di una rapida rimozione, causano una trombocitopenia⁴, trombosi microvascolare e infarti tissutali.

VWD 2M e 2N

Queste due tipologie sono molto più rare delle precedenti poiché dipendono da mutazioni del VWF legate all'inattività dei siti attivi per i leganti alle piastrine e al tessuto connettivo della ECM. Ciò nonostante, le distribuzioni dei mutimeri e i livelli di concentrazioni plastiche del VWF sono nelle norma.

La mutazione nella tipologia 2M è localizzata nel dominio A1 e impedisce il legame con le GPIb delle piastrine. Test di laboratorio portano a diagnosi simili per il tipo 2A ma in questo caso i *large multimers* sono presenti.

La VWD tipo 2N (N sta per "Normandia", luogo dei primi casi) comporta diverse mutazioni del VWF che lo rende inattivo per i siti del fattore VIII, in questo modo la funzione piastrino-dipendente del VWF è presente ma i livelli del fattore VIII sono più bassi del 10%, causando

⁴Per trombopenie si intendono malattie che causano una riduzione di piastrine o trombociti inferiori a 150000 U/mm³

un'emofilia di tipo autosomico. La diagnosi della VWD tipo 2N rispetto all'emofilia A consta nella mancata risposta alla terapia con fattore VIII, dove invece per l'emofilia si ha risposta positiva.

1.4.3 VWD di tipo 3

La VWD di tipo 3 è un disordine recessivo sovrapponibile con la variante di tipo 1, ma data la totale assenza del VWF, presenta dei sintomi talmente severi da rappresentare terapeuticamente una classe distinta.

L'assenza della proteina VWF causa una secondaria deficienza del fattore VIII. I pazienti affetti presentano problemi legati all'adesione piastrinica e nella cascata coagulativa che creano seri problemi emorragici fin dalla nascita. Il fattore VIII è ridotto a livelli della tipologia 2N e i sintomi sono molto simili se non più gravi dell'emofilia A.

1.5 Diagnosi

La VWD presenta uno spettro molto complesso e ampio di varianti. Questo si traduce in una difficile diagnosi per il paziente che presenta i sintomi e una dubbia identificazione, in certi casi, della tipologia di appartenenza.

Un pannello di esperti del National Heart, Lung and Blood Institute (NHLBI) per la valutazione della VWD e di altri disordini della coagulazione, ha raccolto le raccomandazioni cliniche e diagnostiche, basate sull'evidenza, per queste patologie (Nichols, 2007). Inizialmente si effettua un'indagine storica del paziente sugli aspetti chiave clinici per determinare se il paziente stesso ha avuto dei benefici da diagnosi effettuate in precedenza. Dopo un'analisi storica ed uno screening dei segni e sintomi, nel caso fossero comparsi per la prima volta, si passa ad una fase di test di analisi in laboratorio. I test iniziali più utilizzati sono tre:

- misura della concentrazione di antigene del VWF (VWF:AG) presente nel plasma;
- misura dell'attività del cofattore ristocetina (VWF:RCo);
- misura della funzionalità del VWF come vettore per il fattore VIII;

Il VWF:AG è un test immunologico che determina la concentrazione delle proteine del VWF nel plasma e solitamente è basato su metodi di analisi di tipo immuni-assorbenti ELISA (Saggio Immuno-Assorbenti Legato ad un Enzima), che analizza le molecole di antigene catturate da un anticorpo al quale è legato un enzima.

La VWF:RCo è un saggio funzionale del VWF e misura l'abilità di interagire con le normali

piastrine; la ristocetina è un antibiotico che attiva il VWF il quale legandosi alle piastrine crea un coagulo che viene rimosso dalla circolazione. Questo tipo di test non è molto impiegato perché può causare trombocitopenia.

La misura del fattore VIII invece descrive la funzione del VWF come legante e veicolo nel processo di emostasi secondaria, e soprattutto riesce a distinguere la VWD dai casi di emofilia.

I test descritti vengono eseguiti con dei prelievi sanguigni a diversi istanti temporali di interesse, fino ad una durata totale che può arrivare a 24 h, dopo la somministrazione intracutanea di desmopressina (DDAVP), un ormone sintetico che promuove il rilascio del VWF dalle cellule endoteliali. I risultati di tutti i test vengono poi riportati in termini di concertazione di in unità per decilitro [U/dL] o unità internazionali per decilitro [IU/dL] se confrontati con degli standard di riferimento dati dalla World Health Organisation (WHO).

Ulteriori analisi più approfondite vengono effettuate per definire con maggior sicurezza i sottotipi della VWD, e sono basate sulla misurazione sia quantitativa che qualitativa, del VWF e dei suoi multimeri lisi dall'enzima ADAMTS-13. Il test effettuato alle stesse modalità dei precedenti è il binding al collagene (VWF:CB): viene misurata la concentrazione di collagene che si lega nel dominio A3 del VWF. È un test molto importante perché il legame dipende fortemente dalla lunghezza dei multimeri del VWF, infatti il collagene si lega maggiormente con i multimeri di grandi dimensioni rispetto a quelli più leggeri, quindi ci dà informazioni sulle concentrazioni [U/dL] di multimeri a maggior peso molecolare.

Fino ad ora i test elencati presentano carattere quantitativo. Una misura più onerosa che viene eseguita quasi sempre successivamente ai test iniziali, è la valutazione qualitativa dei multimeri tramite un'analisi dei campioni di gel-elettroforesi con la proteina sodio-dodecilsulfato (SDS), utilizzando un anticorpo o una combinazione di anticorpi radiomarcanti e quindi visibili ad una successiva radiografia (Figura 1.4). Si cercherà di limitare l'analisi di elettroforesi solo nei casi in cui la diagnosi è fortemente dubbiosa, essendo un test oneroso, in questi casi sarà possibile distinguere più facilmente la distribuzione e il carattere dei multimeri o la loro mancanza.

Esistono ulteriori metodi per effettuare ed affinare la diagnosi della VWD, ma purtroppo il rilascio del VWF, e quindi l'indagine delle sue disfunzioni, è soggetto ad un ampio spettro di variabilità sia metodologiche, in base al tipo e alle modalità del test effettuato, che dipendenti dal soggetto, come il gruppo sanguigno piuttosto che l'esercizio fisico praticato dal paziente stesso, oppure dal livello di ansia, dalla tipologia di farmaci assunti o dal tipo di interventi chirurgici

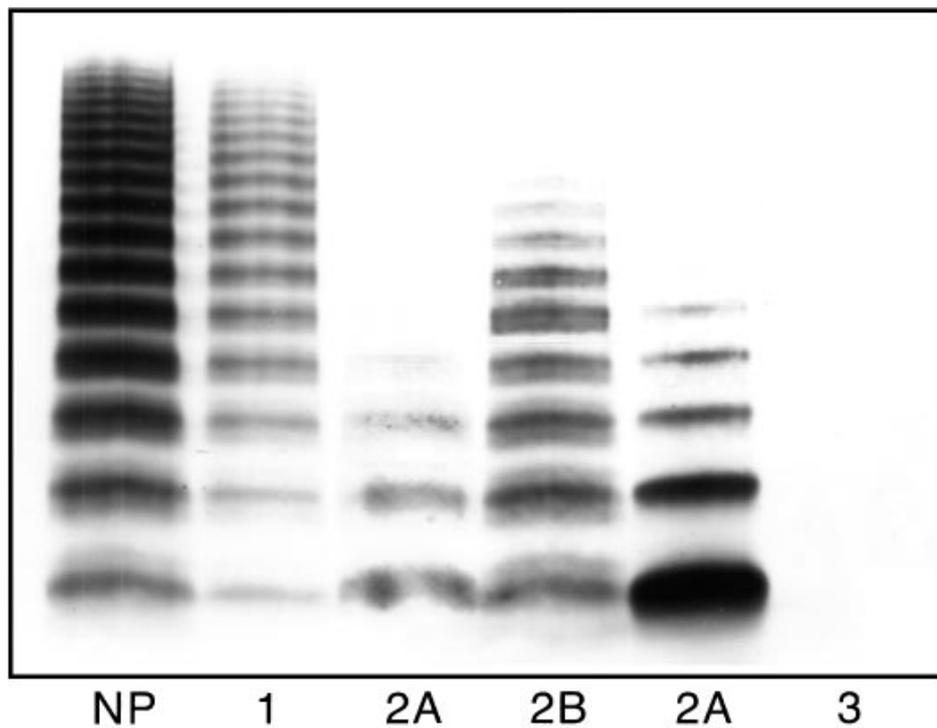


Figura 1.4: Rappresentazione delle diverse distribuzioni multimeriche del VWF sottoposto ad una analisi di gel elettroforesi 1.3% agarosio/SDS. Si notino i diversi pattern rispetto ai pazienti normali (NP), la mancanza dei multimeri ad alto peso molecolare nelle tipologie 2 e la totale assenza nella tipologia 3 (Sadler, 1998).

subiti, una serie di concause che possono aumentare i livelli basali del VWF mascherandone i valori inferiori e anomali.

1.6 Obiettivi dello studio

Si è visto come il VWF sia importante per tutto il sistema coagulativo e come le sue mutazioni e disfunzioni causino una grande varietà di anomalie qualitative e quantitative (Figura 1.4).

Indipendentemente da quale fase del processo sia affetta, queste mutazioni causano degli effetti dominanti e producono distinti fenotipi della VWD. Mutazioni nelle subunità-VWF e successiva non corretta polimerizzazione e rilascio causano difetti quantitativi (tipo 1), mutazioni che interferiscono con le distribuzioni delle unità multimeriche o stabilità delle stesse causano difetti qualitativi (tipo 2), mentre mutazioni che rendono inattivi i siti di legame, o regolano le affinità con le piastrine, causano ulteriori fenotipi che chiaramente spiegano l'importanza biologica della dipendenza delle piastrine e del fattore VIII dal VWF.

Come accennato nel paragrafo §1.1, l'obiettivo principale di questo lavoro di tesi è quello di ottenere attraverso un approccio PSE un modello meccanicistico per lo studio della VWD. L'im-

portanza di avere a disposizione un modello matematico affidabile, basato sulla fisiologia del singolo soggetto, che descriva in modo quantitativo la distribuzione dei multimeri di VWF nel tempo, rappresenta un passo in avanti per: aumentare la comprensione dei complessi fenomeni fisiologici che avvengono nei pazienti affetti da VWD; fornire gli strumenti necessari in modo tale da proporre un protocollo per una diagnosi più sicura e veloce; consentire delle simulazioni in silico per aprire la strada verso una terapia più efficiente.

Le tecniche formali che saranno adottate per affrontare il problema sono discusse nel prossimo capitolo.

Capitolo 2

Strumenti matematici e statistici per lo sviluppo ottimale di un modello dinamico

L'attività di modellazione rappresenta uno strumento indispensabile come supporto alle attività di progettazione, di ottimizzazione e di controllo dei processi chimici. L'obiettivo principale è quello di sviluppare un modello matematico che rappresenti nel modo più fedele il comportamento di un fenomeno deterministico. Lo sviluppo e la verifica di modelli stazionari o dinamici sono aspetti chiave nell'ingegneria di processo e il capitolo seguente descrive la base teorica degli strumenti matematici e statistici utilizzati in questo lavoro di tesi per la fase di modellazione di sistemi fisiologici.

2.1 Introduzione

Durante la costruzione di un modello si cerca di spiegare il fenomeno osservato tramite l'utilizzo di conoscenze a priori, espresse da leggi fisiche, chimiche e biologiche, proponendo diverse formulazioni di modelli possibili. È fondamentale capire quale tra le diverse strutture possibili descriva il sistema di interesse nel modo migliore. Inoltre questi modelli contengono dei parametri con un significato fisico, è necessario sapere se sarà sempre possibile determinarne il loro valore, e con che grado di precisione, da un pool di dati sperimentali.

In questa tesi si descriveranno dei modelli fisiologici la cui caratteristica principale è data dalla complessità (Carson e Cobelli 2001). Questa complessità è dovuta dalle strette interconnessioni che gli elementi di sistemi fisiologici (molecole, cellule, organi e organismo) creano, generando schemi gerarchici affetti da non linearità e variabilità stocastica. Forme di complessità crescente sono presenti quando, all'interno di un organismo, vengono osservati dei meccanismi di controllo che apparentemente non danno alcun tipo di risposta al cambiare delle variabili fisiologiche. Altre caratteristiche dei sistemi fisiologici sono: poca osservabilità poiché è molto difficile ottenere delle misure direttamente *in vivo*, e nella maggior parte dei casi sono disponibili solo delle misure indirette; poca verificabilità, essendo molto difficile avere il pieno controllo di un sistema fisiologico, è necessario operare imponendo dei vincoli specifici di sicurezza e mantenendo

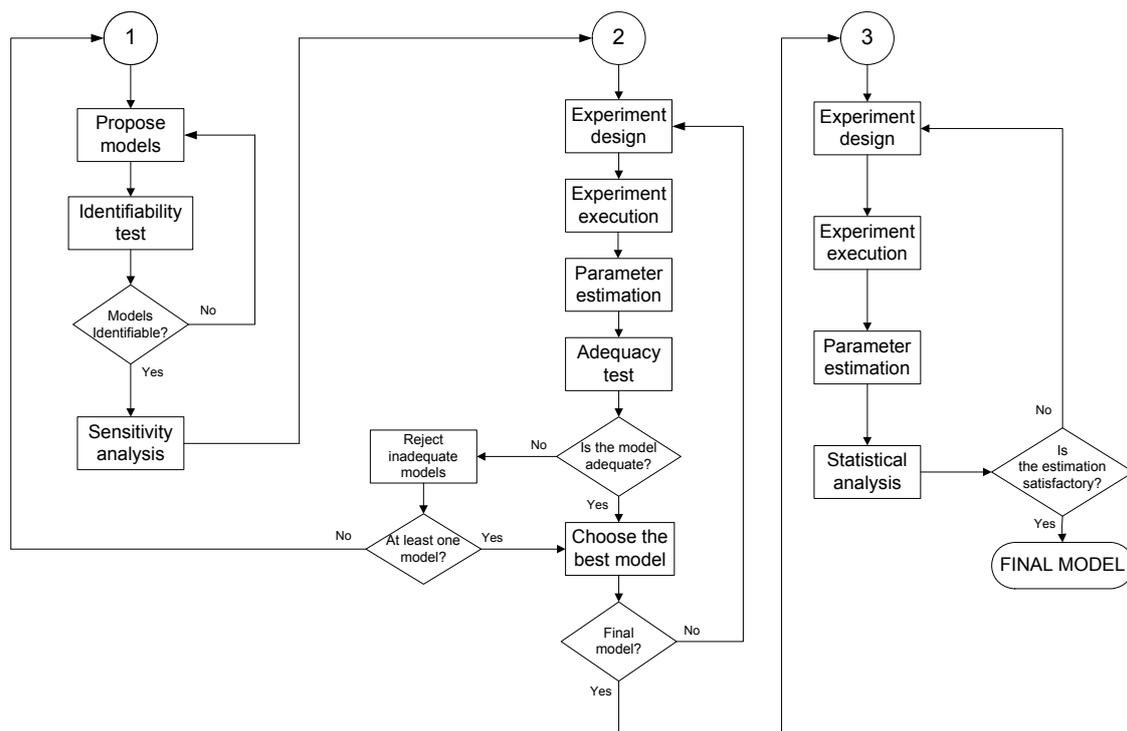


Figura 2.1: Strategia di costruzione di un modello (Galvanin 2010 basata sullo schema di Asprey e Macchietto 2000)

funzionale il comportamento del sistema stesso.

Modelli fisiologici sono stati proposti in campo farmacologico e medico per analizzare e rappresentare il comportamento metabolico di organi e apparati. Essendo una rappresentazione di una realtà complessa questi modelli matematici presentano un certo grado di approssimazione. Il risultato di questa notevole complessità si traduce nel modello con struttura \mathbf{M} distinto da un set di parametri θ . Se il numero di parametri è elevato si può arrivare verso una condizione di sovrapparametrizzazione, la quale solitamente è legata alla non identificabilità del modello.

I metodi seguenti possono essere utilizzati per verificare se sarà possibile stimare il set di parametri del modello da dati sperimentali, partendo dalla conoscenza preliminare limitata del sistema e se ciò sarà possibile, con che grado di precisione.

2.2 Il processo di modellazione

Una procedura sistematica per l'attività di modellazione è stata proposta da Asprey e Macchietto (2000), espressa in Figura 2.1, come supporto allo sviluppo e alla verifica di modelli dinamici di processi lineari e non lineari. Tale procedura è articolata in tre fasi:

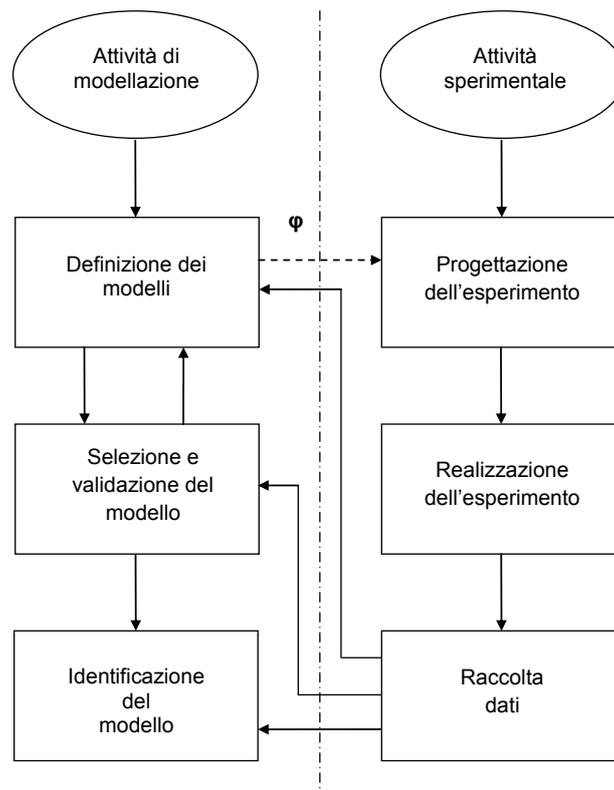


Figura 2.2: Relazione tra attività di modellazione e di sperimentazione; con la freccia tratteggiata è indicato il ponte di collegamento tra le due aree nel caso in cui (MBD_{oE}) il vettore dell'esperimento φ sia ricavato a partire dal modello (Galvanin, 2010).

- I: analisi preliminare che include il test di identificabilità dei modelli proposti;
- II: progettazione ottimale degli esperimenti basata su modello per discriminare i modelli rivali;
- III: progettazione ottimale degli esperimenti basata su modello per aumentare la precisione dei parametri.

Nella fase preliminare vengono proposti diversi modelli ed analizzati tramite test di identificabilità, per verificare la possibilità di stimare sperimentalmente il set di parametri. La seconda fase è caratterizzata da una progettazione di esperimenti al fine di discriminare i diversi modelli rivali. I modelli inadeguati vengono scartati e il modello migliore (o i migliori) viene scelto per la terza fase, nella quale vengono progettati degli esperimenti per migliorare la precisione dei parametri al di sopra di una certa soglia. Si noti come le attività di modellazione e di sperimentazione si completino a vicenda attraverso un'integrazione biunivoca continua, come illustrato in Figura 2.2.

Si consideri la forma generale di un modello deterministico $\mathbf{M}(\boldsymbol{\theta})$ come un insieme di equazioni

differenziali ed algebriche (DAE):

$$\mathbf{M}(\boldsymbol{\theta}) = \begin{cases} \mathbf{f}(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \mathbf{w}, \boldsymbol{\theta}, t) = 0 \\ \mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t)) \end{cases} \quad (2.1)$$

dove $\mathbf{x}(t) \in \mathbb{R}_S^N$ è il vettore delle variabili di stato dipendenti dal tempo, $\mathbf{u}(t) \in \mathbb{R}_U^N$ e $\mathbf{w} \in \mathbb{R}_W^N$ rispettivamente le variabili manipolate dipendenti ed indipendenti dal tempo, $\boldsymbol{\theta} \in \Theta$ il vettore contenente il set di parametri del modello, $\mathbf{y}(t) \in \mathbb{R}_Y^N$ il vettore delle risposte stimate del modello e t il tempo.

L'obiettivo principale che ci si pone è quello di ottenere delle condizioni sperimentali ottimali in modo tale da ottenere dei dati sperimentali sui quali identificare il modello $\mathbf{M}(\boldsymbol{\theta})$, attraverso la stima accurata del set $\boldsymbol{\theta}$. È necessario quindi avere a disposizione un modello tramite il quale si progetteranno degli esperimenti (*model based design of experiment* MBDoe).

Il vettore di progettazione degli esperimenti $\boldsymbol{\varphi}$ è definito nel modo seguente:

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}(\mathbf{u}(t), \mathbf{w}, \mathbf{y}_0, \mathbf{t}_{sp}, \tau) \quad (2.2)$$

dove \mathbf{t}_{sp} è il vettore dei N_{sp} tempi di campionamento, τ la durata dell'esperimento e \mathbf{y}_0 rappresenta il set di condizioni iniziali sulle variabili misurabili.

All'interno di tale vettore vengono inserite tutte le possibili variabili di controllo utilizzabili durante l'esperimento. Tali variabili possono essere modificate sia inizialmente che durante il corso della prova, ottenendo un'ottimizzazione della progettazione, variandone opportunamente il loro valori.

2.3 Analisi preliminare

La validazione strutturale di un modello è di fondamentale importanza per poter affrontare le fasi successive della procedura di modellazione. L'obiettivo della fase preliminare è quello di avere un modello identificabile necessario per poter eseguire una progettazione degli esperimenti basata, appunto, sul modello.

La fase preliminare è caratterizzata dai seguenti punti fondamentali:

- analisi fenomenologica per la definizione di una possibile struttura del modello;
- stima preliminare del set di parametri $\boldsymbol{\theta}_0$ da dati sperimentali;

- test per l'identificabilità del modello.

Come accennato precedentemente, l'obiettivo di questa prima fase è quello di inserire in fase di progettazione degli esperimenti (II fase della procedura) dei modelli identificabili, quindi sarà necessario verificare la possibilità di ottenere l'identificazione parametrica prima di poter effettuare l'esperimento.

Se la struttura del modello non risulta identificabile non si potrà proseguire nelle fasi successive II e III della procedura in Figura 2.1.

2.3.1 Analisi di sensitività

L'analisi di sensitività permette di studiare il comportamento del modello in oggetto, tramite il ripartizionamento qualitativo e quantitativo nelle risposte dei diversi contributi alle variazioni nelle entrate del sistema. Gli obiettivi dell'analisi di sensitività sono: trovare quali parametri del set sono i più influenti nelle risposte del modello e analizzare l'andamento dell'informazione per un dato set di condizioni sperimentali.

Il secondo punto sarà di particolare interesse per la formulazione del concetto di matrice di informazione dinamica spiegato successivamente (vedi §2.3.3).

La valutazione dell'impatto di variazioni nei valori dei parametri stimati $\hat{\theta}$ del modello per il sistema di risposte \hat{y} può essere effettuata studiando i profili temporali della matrice $N_y \times N_\theta$ di sensitività locale:

$$\mathbf{Q}(t) = \begin{bmatrix} q_{1,1}(t) & \cdots & q_{1,N_\theta}(t) \\ \vdots & \ddots & \vdots \\ q_{N_y,1}(t) & \cdots & q_{N_y,N_\theta}(t) \end{bmatrix} = \begin{bmatrix} \frac{\partial \hat{y}_1(t)}{\partial \theta_1} & \cdots & \frac{\partial \hat{y}_1(t)}{\partial \theta_{N_\theta}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}_{N_y}(t)}{\partial \theta_1} & \cdots & \frac{\partial \hat{y}_{N_y}(t)}{\partial \theta_{N_\theta}} \end{bmatrix} \quad (2.3)$$

Ciascun elemento $q_{i,j}(t) = \frac{\partial y_j}{\partial \theta_i} \in \mathbf{Q}(t)$ descrive il comportamento dinamico del sistema parametrico e fornisce un'indicazione per l'allocazione ottimale delle misurazioni sperimentali affinché si possa ottenere il massimo dell'informazione per la stima del parametro θ_i , $i = 1, \dots, N_\theta$ dalla risposta y_j , $j = 1, \dots, N_y$.

Un esempio dei profili temporali degli elementi $q_{i,j}(t)$ della matrice (2.3) è rappresentato in Figura 2.3, si noti come i parametri θ_1 e θ_4 abbiano un effetto più marcato sulla risposta del modello rispetto a θ_2 e θ_3 .

La matrice (2.3) è valutata per la r -esima risposta misurata a ciascuno tempo di campionamento,

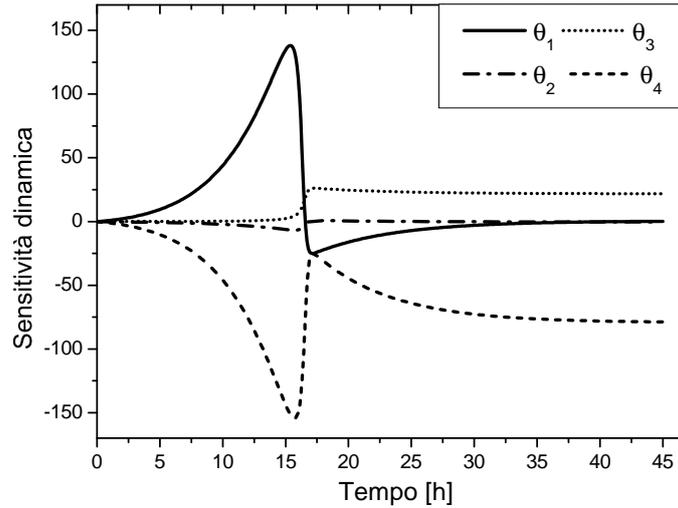


Figura 2.3: Esempio di profili di sensibilità dinamica rispetto al set di parametri di un modello. In figura sono rappresentati gli elementi di (2.3) per una risposta $\mathbf{y}(t)$ del modello.

attraverso la matrice di sensibilità \mathbf{Q}_r di dimensioni $n_{sp} \times N_\theta$:

$$\mathbf{Q}_r = \begin{bmatrix} \left. \frac{\partial y_r}{\partial \theta_1} \right|_{t_1} & \cdots & \left. \frac{\partial y_r}{\partial \theta_{N_\theta}} \right|_{t_1} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial y_r}{\partial \theta_1} \right|_{t_{n_{sp}}} & \cdots & \left. \frac{\partial y_r}{\partial \theta_{N_\theta}} \right|_{t_{n_{sp}}} \end{bmatrix} \quad (2.4)$$

essa descrive la sensibilità della risposta r -esima al set di parametri del modello, in altre parole rappresenta in forma discreta gli elementi della riga i -esima di $\mathbf{Q}(t)$ in (2.3).

2.3.2 Test di identificabilità strutturale

I test di indentificabilità a priori vengono contraddistinti in due tipologie: test di identificabilità globale e test di identificabilità locale.

I primi hanno lo scopo di dimostrare la possibilità di stimare i parametri come proprietà del modello, indipendentemente dai dati misurati e dall'incertezza connessa alla misurazione; l'identificabilità strutturale di un modello può essere (Cobelli 2007):

- globale, o univoca, se è verificata la seguente relazione: $\mathbf{M}(\boldsymbol{\theta}) = \mathbf{M}(\boldsymbol{\theta}^*)$ se e solo se $\boldsymbol{\theta} = \boldsymbol{\theta}^*$, tutti i parametri del modello risultano univocamente identificabili;
- locale, o non univoca, se $\mathbf{M}(\boldsymbol{\theta}) = \mathbf{M}(\boldsymbol{\theta}^*)$ per un numero limitato di vettori $\boldsymbol{\theta}^*$, cioè tutti i parametri sono identificabili con l'eccezione di uno o più parametri non univocamente identificabili, ma per i quali si ha un numero di soluzioni finito;
- non-identificabile, $\mathbf{M}(\boldsymbol{\theta}) = \mathbf{M}(\boldsymbol{\theta}^*)$ per ogni θ_i tale per cui almeno un parametro non è

identificabile.

I test d'identificabilità locale hanno il proposito di valutare univocamente un set di parametri con un numero sufficiente di punti sperimentali, limitando l'analisi in una parte circoscritta dello spazio delle soluzioni dei parametri considerati. Essi sono test numerici effettuati in un intorno limitato del set $\boldsymbol{\theta}$ e servono per valutare l'esistenza di ulteriori set parametrici, nell'intorno fissato, in grado di produrre la stessa risposta del modello generata da $\boldsymbol{\theta}$, per determinati valori del vettore $\boldsymbol{\varphi}$.

Una procedura per l'analisi di identificabilità strutturale di un modello è quella di accoppiare un test di identificabilità strutturale locale (SLI) e un test di identificabilità globale (SGI). Partendo da una data struttura $\mathbf{M}(\boldsymbol{\theta})$ di un modello e la conoscenza a priori del set parametrico $\boldsymbol{\theta}_0$ come guess iniziale, e una matrice di varianza-covarianza preliminare dei parametri $\boldsymbol{\Sigma}_0$, si definisce la variabilità attesa nel dominio del set parametrico.

La matrice di varianza-covarianza \mathbf{V}_θ del set dei parametri del modello $\boldsymbol{\theta}$, è costituita nel modo seguente:

$$\mathbf{V}_\theta = \boldsymbol{\Sigma}_\theta = \{v_{ij}\} = \left[\sum_{r=1}^{N_y} \sum_{s=1}^{N_y} \tilde{\sigma}_{r,s} \mathbf{Q}_r^T \mathbf{Q}_s \right]^{-1} \quad (2.5)$$

dove $\tilde{\sigma}_{r,s}$ è la varianza-covarianza per le r,s -esime risposte sulle misure sperimentali, e $\mathbf{Q}_{r,s}$ la matrice di sensitività dinamica (2.3) per le r,s -esime risposte del modello.

Il test SLI proposto da Shaw (1999) è basato sulla determinazione del rango della matrice di stimabilità \mathbf{P}_E , di dimensioni $N_y N_{sp} \times N_\theta$:

$$\mathbf{P}_E = \begin{bmatrix} \mathbf{S}_y(t_0) \\ \vdots \\ \mathbf{S}_y(t_{N_{sp}}) \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial \theta_1} \Big|_{t_0} & \cdots & \frac{\partial y_1}{\partial \theta_{N_\theta}} \Big|_{t_0} \\ \cdots & \ddots & \cdots \\ \frac{\partial y_{N_y}}{\partial \theta_1} \Big|_{t_0} & \cdots & \frac{\partial y_{N_y}}{\partial \theta_{N_\theta}} \Big|_{t_0} \\ \vdots & & \vdots \\ \frac{\partial y_1}{\partial \theta_1} \Big|_{t_{N_{sp}}} & \cdots & \frac{\partial y_1}{\partial \theta_{N_\theta}} \Big|_{t_{N_{sp}}} \\ \cdots & \ddots & \cdots \\ \frac{\partial y_{N_y}}{\partial \theta_1} \Big|_{t_{N_{sp}}} & \cdots & \frac{\partial y_{N_y}}{\partial \theta_{N_\theta}} \Big|_{t_{N_{sp}}} \end{bmatrix} \quad (2.6)$$

dove $\mathbf{S}_y(t_k)$ è la $N_y \times N_\theta$ matrice di sensitività valutata al tempo di campionamento t_k ; se il rango della matrice $\text{rank}(\mathbf{P}_E) < N_\theta$, i coefficienti di sensitività non sono indipendenti e non sarà possibile stimare l'intero set dalla tipologia di dati disponibili.

Una formulazione alternativa per il test SLI in grado di fornire un'indicazione preliminare quantitativa sulla difficoltà nella stima dei parametri, è definita valutando la correlazione c_{ij} delle colonne della matrice \mathbf{P}_E :

$$c_{ij} = \frac{v_{ij}}{\sqrt{v_{ii} \cdot v_{jj}}}, \quad \text{con} \quad v_{ij} = E [(p_{Ei} - \eta_i) (p_{Ej} - \eta_j)] \quad (2.7)$$

dove rispettivamente: v_{ij} è la covarianza del parametro θ_i rispetto al θ_j mentre $v_{ii} = \sigma_i$ la varianza del parametro θ_i , p_{Ei} elemento dell' i -esima colonna di \mathbf{P}_E , E indica il valore atteso e η_i il valor medio degli elementi dell' i -esima colonna di \mathbf{P}_E .

Il calcolo di (2.7) produce una matrice di correlazione \mathbf{C} simmetrica $N_\theta \times N_\theta$.

$$\mathbf{C} = \{c_{ij}\} = \begin{bmatrix} \frac{v_{1,1}}{\sigma_{\theta_1} \sigma_{\theta_1}} & \dots & \frac{v_{1,N_\theta}}{\sigma_{\theta_1} \sigma_{\theta_{N_\theta}}} \\ \vdots & \ddots & \vdots \\ \frac{v_{N_\theta,1}}{\sigma_{\theta_{N_\theta}} \sigma_{\theta_1}} & \dots & \frac{v_{N_\theta,N_\theta}}{\sigma_{\theta_{N_\theta}} \sigma_{\theta_{N_\theta}}} \end{bmatrix} \quad (2.8)$$

Se c_{ij} è diverso da ± 1 il modello è localmente identificabile, se $c_{ij} = \pm 1$ la coppia di parametri $\theta_i \theta_j$ sarà completamente correlata (1) o anticorrelata (-1). In ogni caso risulterà che se le correlazioni assumono valori superiori a 0.99, il sistema diventa non identificabile e non c'è garanzia sulla stima dei parametri, questo può derivare dalla sovrapparametrizzazione del modello, dal numero limitato di dati disponibili o dall'errore di misura sui dati.

Si può notare che al fine del calcolo della correlazione tra le colonne di \mathbf{P}_E le formulazioni seguenti per la matrice di stimabilità risultano equivalenti, da (2.3), (2.4) e (2.6):

$$\mathbf{P}_E = \begin{bmatrix} \mathbf{S}_y(t_0) \\ \vdots \\ \mathbf{S}_y(t_{N_{sp}}) \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_r(\theta_1) \\ \vdots \\ \mathbf{Q}_r(\theta_{N_\theta}) \end{bmatrix}. \quad (2.9)$$

Un esempio di test SGI basato sull'ottimizzazione di identificabilità globale viene proposto da Asprey e Macchietto (2000), definisce il modello, con struttura $\mathbf{M}(\boldsymbol{\theta})$ ed una risposta \mathbf{y} , globalmente identificabile, se per qualsiasi due set parametrici $\boldsymbol{\theta}$ e $\boldsymbol{\theta}^*$, ad un orizzonte temporale $t \in [0, \tau]$, per il vettore delle variabili di controllo \mathbf{u}_0 e le stesse condizioni iniziali \mathbf{y}_0 , la distanza

Φ^J tra i due vettori di parametri che danno la stessa risposta del modello è:

$$\Phi^J = \max_{\boldsymbol{\theta}, \boldsymbol{\theta}^*} (\boldsymbol{\theta} - \boldsymbol{\theta}^*)^T \mathbf{W}_\theta (\boldsymbol{\theta} - \boldsymbol{\theta}^*) \leq \varepsilon_\theta \quad (2.10a)$$

soggetto a

$$\int_0^\tau (\mathbf{y}(\mathbf{u}_0, \boldsymbol{\theta}) - \mathbf{y}(\mathbf{u}_0, \boldsymbol{\theta}^*))^T \mathbf{W}_y (\mathbf{u}_0, \boldsymbol{\theta}) - \mathbf{y}(\mathbf{u}_0, \boldsymbol{\theta}^*) dt \leq \varepsilon_y \quad (2.10b)$$

$$\mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{u}_0, \boldsymbol{\theta}, t) = 0 \quad (2.10c)$$

dove \mathbf{W}_θ e \mathbf{W}_y sono due matrici peso appropriate, ed ε_θ e ε_y sono due numeri arbitrariamente piccoli.

Il test SGI implica la diretta valutazione numerica di identificabilità nell'intero dominio di variabilità ($\mathbf{M}(\boldsymbol{\theta}) = \mathbf{M}(\boldsymbol{\theta}^*)$ se e solo se $\boldsymbol{\theta} = \boldsymbol{\theta}^*$) del set di parametri $\boldsymbol{\theta}$ determinato per $\boldsymbol{\varphi}_0$ nel test SLI. In altre parole il modello può essere identificabile strutturalmente se la distanza tra i due vettori parametrici che danno la stessa risposta è arbitrariamente piccola.

2.3.3 Analisi d'informazione

Avendo a che fare con sistemi complessi multi-input multi-output (MIMO) l'analisi di sensitività conterrà un gran numero di elementi. Il capire come possa influire l'incertezza di una stima parametrica iniziale sulle risposte date dal sistema modellato e, di conseguenza, sulla quantità di informazioni generate dall'esperimento, richiede l'elaborazione di una notevole quantità di dati. Per ovviare a questa problematica e quindi quantificare l'insieme di informazione, si utilizza un modo più diretto introducendo il concetto di matrice di informazione dinamica. Il concetto di matrice di informazione deriva dalla matrice di informazione di Fisher FIM $\mathbf{I}(\boldsymbol{\theta})$, espresso in Soderstrom and Stoica (1977), la quale rappresenta un modo per misurare la quantità di informazione che una variabile osservata random y supporta, circa un vettore di parametri sconosciuto $\boldsymbol{\theta}$. Per sistemi MIMO la matrice di informazione diventa:

$$\mathbf{H}(\boldsymbol{\theta}) = \sum_{r=1}^{N_y} \sum_{s=1}^{N_y} \tilde{\sigma}_{r,s} \mathbf{Q}_r^T \mathbf{Q}_s \quad (2.11)$$

dove $\tilde{\sigma}_{r,s}$ è l'elemento r,s -esimo dell'inverso della matrice di varianza covarianza $\boldsymbol{\Sigma}_y$ sulle misure sperimentali, e $\mathbf{Q}_{r,s}$ la matrice di sensitività dinamica (2.3) per le r,s -esime risposte del modello.

La generica forma della matrice Σ_y assume la struttura:

$$\Sigma_y = \text{cov}(\mathbf{y}_r, \mathbf{y}_s) = \begin{bmatrix} \sigma_{y_1, y_1}^2 & \sigma_{y_1, y_2}^2 & \cdots & \sigma_{y_1, y_{N_s}}^2 \\ \sigma_{y_2, y_1}^2 & \sigma_{y_2, y_2}^2 & \cdots & \sigma_{y_2, y_{N_s}}^2 \\ \vdots & & \ddots & \vdots \\ \sigma_{y_{N_r}, y_1}^2 & \sigma_{y_{N_r}, y_2}^2 & \cdots & \sigma_{y_{N_r}, y_{N_s}}^2 \end{bmatrix} \quad (2.12)$$

dove imporre che gli elementi extradiagonali di Σ_y siano nulli, equivale ad assumere che all'interno dello stesso esperimento, le misure siano completamente indipendenti. Un modo per riassumere tale informazione (2.11) può essere dato dalla sua traccia $\text{tr}(\mathbf{H}(\boldsymbol{\theta}))$ la quale appunto è data dalla somma dei termini diagonali.

Si noti come la definizione di matrice di varianza-covarianza \mathbf{V}_θ (2.5) sia l'inverso della matrice di informazione $\mathbf{H}(\boldsymbol{\theta})$ (2.11):

$$\mathbf{V}_\theta = [\mathbf{H}(\boldsymbol{\theta})]^{-1} = \left[\sum_{r=1}^{N_y} \sum_{s=1}^{N_y} \tilde{\sigma}_{r,s} \mathbf{Q}_r^T \mathbf{Q}_s \right]^{-1} \quad (2.13)$$

Qualora il modello non superi il test SLI (2.6) significa che gli elementi della matrice di sensibilità tendono a zero, la matrice di informazione diventa singolare.

Esistono diversi modi per rimediare a tale problema, modificando la valutazione dell'informazione attesa agendo sugli elementi della matrice (2.3), il più importante è operando una riparametrizzazione del set di parametri $\boldsymbol{\theta}$ (vedi §2.5).

2.4 Stima dei parametri

Qualora si avranno a disposizione dei dati sperimentali sulle risposte del modello, si procederà con la stima dei parametri, cioè si dovranno individuare i valori da attribuire ai parametri del modello per regredire nel modo ottimale i dati acquisiti sulle risposte. Gli obiettivi della stima parametrica sono:

- massimizzare la capacità del modello di predire la risposta misurata, provvedendo un fitting soddisfacente dei dati misurati;
- realizzare una stima accurata dei parametri ottenendo valori aventi determinate statistiche.

Uno stimatore può essere definita statisticamente, come funzione dei dati osservabili \mathbf{y} , nella forma:

$$\hat{\boldsymbol{\theta}} = \Phi^{PE}(\mathbf{y}) : \mathbb{R}^{N_y} \rightarrow \mathbb{R}^{N_\theta} \quad (2.14)$$

dove $\hat{\boldsymbol{\theta}}$ rappresenta il vettore della stima dei parametri, cioè il risultato in uscita dalla stima. Una stima efficiente dev'essere accurata e precisa, ovvero: si devono ottenere dei valori in uscita che siano i più vicini possibili ai parametri veri $\boldsymbol{\theta}$, del modello che descrive il sistema in modo attendibile; i valori stimati devono presentare una dispersione minima, cioè una minima varianza (precisione).

Il risultato ideale di una stima dovrebbe essere un vettore avente la minima varianza ed allo stesso tempo che fornisca la minima deviazione tra risposta predetta $\hat{\mathbf{y}}$ e quella misurata \mathbf{y} . Questo può essere ottenuto minimizzando gli elementi della matrice $N_y \times n_{sp}$ dei residui assoluti \mathbf{r} , dove ciascun elemento è descritto:

$$r_{ij} = \|y_i(t_j) - \hat{y}_i(t_j)\| \quad i = \dots N_y, j = \dots n_{sp} \quad (2.15)$$

Per l'attività di stima dei parametri di modelli dinamici, possono essere utilizzate diverse tecniche con diverse funzioni obiettivo da minimizzare (Bard,1974). L'approccio più utilizzato è quello di massima verosimiglianza. Come risultati si hanno il vettore di stima dei parametri $\hat{\boldsymbol{\theta}}$ del modello ed una distribuzione di probabilità legata alla stima effettuata. Sarà quindi possibile estrarre delle informazioni utili valutando la matrice di varianza-covarianza dei parametri. La funzione obiettivo utilizzata nella ML è:

$$\begin{aligned} \Phi^{ML}(\mathbf{y}, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_N) &= L(\boldsymbol{\theta}, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_N) = \\ &= 2\pi^{\left(\frac{N_y}{2}\right)} \prod_{i=1}^N |\boldsymbol{\Sigma}_i|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^N [(\mathbf{y}_i - \hat{\mathbf{y}}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{y}_i - \hat{\mathbf{y}}_i)] \right\} \end{aligned} \quad (2.16)$$

dove $\boldsymbol{\Sigma}_i$ è la matrice di varianza-covarianza degli errori delle misure nell' i -esimo esperimento. Quando si effettua una stima parametrica basata su ML, per risolvere il problema è necessaria la conoscenza dei seguenti elementi:

- il modello nella forma (2.1) e le condizioni iniziali \mathbf{y}_0 ;
- le stime iniziali dei parametri $\hat{\boldsymbol{\theta}}_0$, descritte eventualmente da delle statistiche a priori che identificano l'incertezza preliminare dello spazio delle soluzioni;
- il set di vettori di design che definiscono le configurazioni dell' N -esimo esperimento

condotto;

- il set \mathbf{y} di punti sperimentali delle risposte di ciascun esperimento;
- l'informazione sul sistema di misurazione, espresso tramite $\Sigma_{\mathbf{y}}$; se essa non è disponibile, sarà consigliabile utilizzare l'utilizzo di un modello di varianza del tipo

$$\sigma_{y_i}^2 = \omega_j^2 (\hat{y}_j^2)^{\gamma_j} \quad (2.17)$$

dove \hat{y}_j è la j -esima risposta predetta del modello, ω_j è la deviazione standard della j -esima risposta misurata dove γ_j è il fattore eteroschedastico. I parametri ω_j e γ_j vengono scelti in base al tipo di modello definito in Tabella 2.1.

I risultati in uscita dalla stima parametrica sono: il vettore dei parametri stimati $\hat{\boldsymbol{\theta}}$, alcune statistiche a-posteriori che definiscono gli intervalli di confidenza delle stime e se richiesto, i parametri stimati per il modello di varianza.

Tabella 2.1: *Modelli di varianza dei parametri.*

Modello di varianza atteso	γ	ω
Varianza costante	0	Fissato a-priori o stimato
Varianza relativa costante	1	Fissato a-priori o stimato
Eteroschedastico	Fissato a priori o stimato	Fissato a-priori o stimato

2.4.1 Analisi della stima dei parametri

Dal punto di vista numerico la fase di stima dei parametri è un problema di integrazione non-lineare per la risoluzione del sistema di equazioni differenziali del modello $\mathbf{M}(\boldsymbol{\theta})$, assieme ad una robusta ottimizzazione per raggiungere un minimo locale durante la procedura di minimizzazione della funzione obiettivo Φ^{PE} . Sarà molto importante quindi affiancare delle verifiche statistiche per controllare la qualità delle stime.

Una stima parametrica deve avere due principali caratteristiche (Emery, 2001): accuratezza, il valore dei parametri deve catturare l'informazione contenuta nelle misure sperimentali trascurando l'effetto del rumore e dei disturbi; precisione, il set di parametri dev'essere confinato in una regione di confidenza ristretta. La precisione delle stime è strettamente correlata alla regione di incertezza descritta dalla matrice di varianza-covarianza $\mathbf{V}_{\boldsymbol{\theta}}$ (2.5).

L'intervallo di confidenza delle stime è di notevole supporto per capire se i parametri sono sti-

mati nel migliore dei modi, e solitamente vengono valutati dalla matrice di varianza-covarianza nel modo seguente:

$$\kappa_i = t \left(\frac{1 - \alpha}{2}, n_{sp}NM - N_\theta \right) \sqrt{v_{ii}} \quad i = 1 \dots N - \theta \quad (2.18)$$

dove t è il limite massimo per il valore critico $(1 - \alpha)/2$ di una distribuzione- t con $n_{sp}NM - N_\theta$ gradi di libertà. Approssimativamente per un valore critico del $(1 - \alpha)/2 = 95\%$ il livello di confidenza è:

$$\kappa_i^{95\%} = 2\sqrt{v_{ii}} \quad i = 1 \dots N - \theta \quad (2.19)$$

e l'intervallo di confidenza è direttamente valutato dalla diagonale di \mathbf{V}_θ . Se i parametri sono assunti come normalmente distribuiti, sarà possibile effettuare il t -test una volta nota la matrice \mathbf{V}_θ :

$$t = \frac{\hat{\theta}_i}{\sqrt{v_{ii}}} \quad i = 1 \dots N - \theta \quad (2.20)$$

dove v_{ii} sono gli elementi diagonali di \mathbf{V}_θ .

I valori di t sono un indicatore per misurare la confidenza dei parametri del modello e, durante il test, dovranno essere comparati ad un valore di riferimento, dato dalla distribuzione t -student per $n_{sp}NM - N_\theta$ gradi di libertà. Se il valore ottenuto della stima sarà superiore a quello di riferimento, la stima è soddisfacente. Valori alti di t significano che i parametri sono stati stimati con un intervallo molto elevato.

Per verificare l'appropriata minimizzazione dei residui, si utilizza il test sul χ^2 (*lack-of-fit-test*), il quale fornisce la somma pesata dei residui (sum of weighted residual SWR):

$$SWR = \sum_{i=1}^N [(y_i - \hat{y}_i)\boldsymbol{\Sigma}_i^{-1}(y_i - \hat{y}_i)]. \quad (2.21)$$

Anche in questo caso il valore in uscita del SWR verrà confrontato con un valore di riferimento di χ^2 dato da una distribuzione con $n_{sp}NM - N_\theta$ gradi di libertà, e se $SWR \leq \chi_{RIF}$ allora il fitting dei dati sperimentali è efficiente e il modello può essere considerato una rappresentazione affidabile del sistema fisico.

2.5 Riparametrizzazione

Quando si effettua una riparametrizzazione viene definito un set di equazioni che siano una combinazione lineare di due o più modelli, in modo tale da creare o modificare il set di parametri

da stimare. Per l' i -esimo parametro del modello sarà:

$$\theta'_i = f_{\theta_i}(\theta_1, \dots, \theta_{N_\theta}) \quad (2.22)$$

ed il nuovo coefficiente della sensitività sarà correlato agli elementi in (2.3) attraverso la seguente relazione:

$$q_i = \frac{\partial \hat{y}}{\partial \theta'_i} = \frac{\partial f_{\theta_i}}{\partial \theta'_i} + \frac{\partial}{\partial \theta'_i} f_{\theta_i} \cdot \frac{\partial \hat{y}}{\partial \theta'_i} = \frac{\partial f_{\theta_i}}{\partial \theta'_i} + \frac{\partial}{\partial \theta'_i} f_{\theta_i} \cdot q'_i \quad (2.23)$$

I benefici, dovuti dall'adozione di un approccio di questo tipo, vengono dalla riformulazione di una nuova matrice di informazione dinamica la quale utilizzerà le sensitività del nuovo sistema riparametrizzato.

Un'elevata correlazione presente tra due o più parametri del modello si riflette in un problema di identificabilità parametrica, spesso incontrato quando il modello è sovrapparametrizzato. Si dovrà adottare un nuovo metodo che sciolga questa complicità senza alterare la struttura del modello, e a partire dai parametri di quest'ultimo, creerà un nuovo set parametrico al fine di generare una nuova matrice di informazione con un maggiore contenuto informativo.

La nuova matrice di informazione sarà basata sul calcolo di sensitività rispetto al nuovo set di parametri θ' , il quale potrà essere ottenuto tramite:

$$\theta' = h(\theta) \quad (2.24)$$

dove la $h(\theta)$ è una funzione invertibile in grado di creare un nuovo vettore parametrico che abbia significato fisico e che sia capace di non eliminare parte dell'informazione presente nel modello. Se si è interessati al valore dei parametri iniziali è necessario imporre l'invertibilità di $h(\theta)$ in modo tale da stimare indirettamente θ una volta stimato θ' attraverso:

$$\theta = h^{-1}(\theta') \quad (2.25)$$

La scelta di $h(\theta)$ dipende dal tipo di modello oggetto di studio e dalle problematiche specifiche emerse da una fase preliminare insoddisfacente effettuata con criteri standard. Attualmente non esiste un criterio univoco di scelta che permetta di associare ad una determinata causa di non identificabilità, una specifica forma ottimale di $h(\theta)$, e l'effettiva validità del metodo sarà testata

con procedure *trial and error*.

La nuova matrice di sensitività dinamica, una volta scelta la forma di $h(\boldsymbol{\theta})$, avrà la forma seguente:

$$\mathbf{Q}'_r = \begin{bmatrix} \left. \frac{\partial y_r}{\partial \theta'_1} \right|_{t_1} & \cdots & \left. \frac{\partial y_r}{\partial \theta'_{N\theta}} \right|_{t_1} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial y_r}{\partial \theta'_1} \right|_{t_{nsp}} & \cdots & \left. \frac{\partial y_r}{\partial \theta'_{N\theta}} \right|_{t_{nsp}} \end{bmatrix} \quad (2.26)$$

Sostituendo la (2.26) nella (2.11) sarà possibile ricavare la nuova matrice di informazione riparametrizzata $\mathbf{H}(\boldsymbol{\theta}')$.

Un criterio per riparametrizzare il nuovo set parametrico viene proposto in De Luca (2012), dove il nuovo set $\boldsymbol{\theta}'$ viene scritto come il prodotto tra una matrice \mathbf{A} , chiamata matrice di riparametrizzazione, e il set $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}' = h(\boldsymbol{\theta}) = \mathbf{A} \cdot \boldsymbol{\theta} \quad (2.27)$$

Il problema della scelta dei parametri $\boldsymbol{\theta}'$ si sposta sulla scelta di un'opportuna matrice \mathbf{A} ma questo porterà a ridefinire l'intera procedura di Figura 2.1. Poiché non ci sarà sempre una relazione univoca tra la stima del nuovo set $\boldsymbol{\theta}'$ ed il vecchio $\boldsymbol{\theta}$, ciò porterà il sorgere di differenti problematiche le quali saranno risolte ridefinendo l'intera procedura di design.

2.6 Progettazione degli esperimenti basata sul modello

Generalmente la progettazione ottimale degli esperimenti dipende fortemente dall'obiettivo finale dell'esperimento, che sia quello di ottenere la minimizzazione dell'errore sulla stima parametrica o il superamento delle incongruenze tra il modello e il sistema reale, (Figura 2.1), dove il vettore $\boldsymbol{\varphi}$ definito in (2.2), potrà assumere valori differenti.

L'ottimizzazione di $\boldsymbol{\varphi}$ è ottenuta attraverso un compromesso tra massimizzazione dell'informazione sulla stima parametrica preliminare e minimizzazione delle risorse utilizzate per condurre le prove sperimentali, tempo, persone, denaro. L'MBDoE richiede la conoscenza dei seguenti dati preliminari:

- stima parametrica iniziale $\boldsymbol{\theta}_0$ e informazioni sull'incertezza parametrica iniziale;
- scelta preliminare di un opportuno criterio di progettazione;
- definizione delle variabili misurate \mathbf{y} e dell'errore sulle misure sperimentali;
- definizione del generico vettore dell'esperimento $\boldsymbol{\varphi}$ in modo da stabilire il campo di

variabilità su cui poter effettuare la progettazione.

La progettazione ottimale di un esperimento consiste nel modificare il vettore $\boldsymbol{\varphi}$ in maniera tale da ottimizzare una specifica funzione obiettivo, e generalmente (Zullo, 1991) essa è costituita dalla matrice di informazione $\mathbf{H}(\boldsymbol{\theta}, \boldsymbol{\varphi})$ definita in (2.11).

Qualora si effettuano N_{exp} esperimenti in sequenza, l'informazione può essere espressa come:

$$\mathbf{H}_{\theta}(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \sum_{j=1}^{N_{exp}} \sum_{r=1}^{N_y} \sum_{s=1}^{N_y} \tilde{\sigma}_{r,s} \mathbf{Q}_r^T \mathbf{Q}_s + \mathbf{H}_{\theta}^0 \quad (2.28)$$

dove \mathbf{H}_{θ}^0 rappresenta la matrice di informazione iniziale, spesso trascurabile in quanto all'inizio della procedura l'incertezza sui parametri è molto alta. La massimizzazione di una metrica di \mathbf{H}_{θ} si traduce in un problema in termini di minimizzazione della metrica della matrice di varianza-covarianza dei parametri \mathbf{V}_{θ} (2.5), dalla quale si ricava la matrice $\boldsymbol{\Sigma}_{\theta}$ di varianza-covarianza sui parametri (2.5).

I criteri di progettazione più noti sono i cosiddetti criteri classici alfabetici:

- criterio *D-ottimale*: in questo criterio la progettazione mira alla massimizzazione del determinante della matrice di informazione (o minimizzazione del determinante della matrice di varianza-covarianza), $\psi(\mathbf{H}_{\theta}^{-1}) = \psi(\mathbf{V}_{\theta}) = \det(\mathbf{V}_{\theta})$;
- criterio *E-ottimale*: utilizzando questo criterio la progettazione mira alla massimizzazione del più piccolo autovalore della matrice di informazione (minimizzazione del più grande autovalore della matrice di varianza-covarianza), $\psi(\mathbf{H}_{\theta}^{-1}) = \psi(\mathbf{V}_{\theta}) = \arg \min \lambda_1(\mathbf{V}_{\theta})$;
- criterio *A-ottimale*: in questo criterio la progettazione mira alla massimizzazione della traccia della matrice di informazione (minimizzazione della traccia della matrice di varianza-covarianza) $\psi(\mathbf{H}_{\theta}^{-1}) = \psi(\mathbf{V}_{\theta}) = \max \text{tr}(\mathbf{V}_{\theta})$.

La differenza tra i diversi criteri può essere compresa usando un'interpretazione geometrica del problema: l'applicazione al set parametrico della matrice di varianza-covarianza genera nello spazio dei parametri un iperellissoide i cui assi sono coincidenti con gli autovettori di \mathbf{V}_{θ} e il cui volume è uguale al determinante di \mathbf{V}_{θ} .

Dalla Figura 2.4 si osserva che il criterio *D-ottimale* mira a ridurre il volume dell'iperellissoide, quello *E-ottimale* minimizza le dimensioni dell'asse maggiore e quello *A-ottimale* cerca di ridurre le dimensioni delle diagonali del solido in cui è inscritto l'iperellissoide. In genere il criterio più utilizzato è il *D-ottimale* poiché ha una facile interpretazione geometrica e rimane invariato rispetto a particolari trasformazioni, come il rescaling dei parametri, oppure alla capacità di

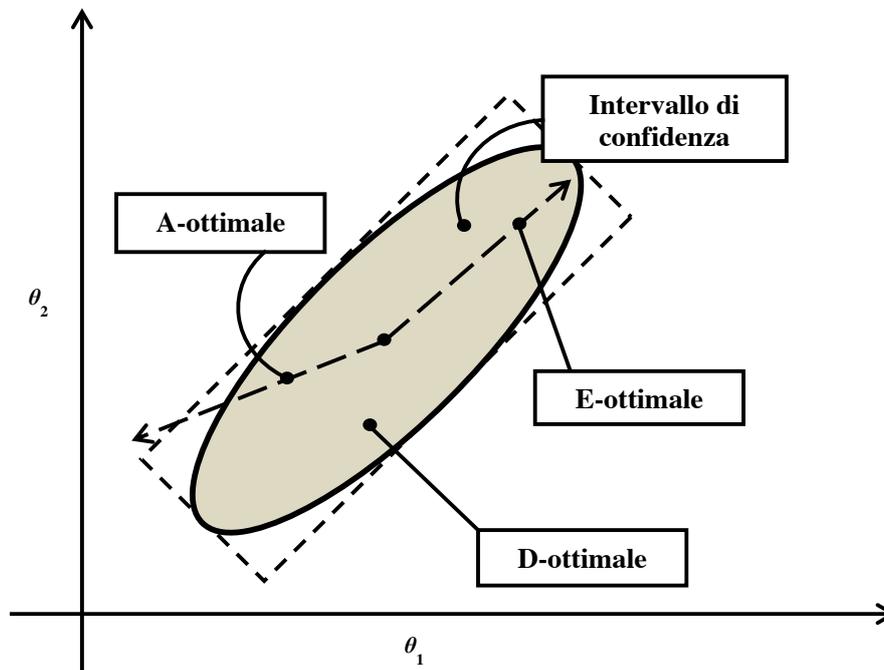


Figura 2.4: Interpretazione geometrica dei criteri standard di design degli esperimenti; l'area in grigio rappresenta l'intervallo di confidenza dei parametri, in genere al 90% o al 95% (Bard e Wolfowitz, 1965)

progettare esperimenti ottimali alterando un numero ristretto di condizioni operative.

2.7 Discriminazione tra diversi modelli proposti

Nella maggior parte dei casi in fase di costruzione di un modello, è possibile che più di un modello venga proposto per descrivere lo stesso fenomeno osservato. Se tutti i modelli proposti superano il test di identificabilità e per ciascuno di essi si riuscirà ad ottenere il sistema di $\mathbf{M}(\boldsymbol{\theta})$ e del suo set parametrico $\boldsymbol{\theta}$, sarà necessario scegliere quale tra i diversi modelli rappresenterà nel modo più adeguato il fenomeno. Emergerà quindi un problema di discriminazione tra N_M modelli, poiché non si saprà con certezza quale sia il migliore.

La strategia di base è quella di attuare un metodo sistematico per individuare le condizioni sperimentali che validificano un modello tra tutti gli altri. Sarà necessario progettare un vettore $\boldsymbol{\varphi}$ di condizioni sperimentali in modo tale da massimizzare una funzione obiettivo e ciò significa che per ogni esperimento il modello migliore provvederà la più accurata predizione delle misure osservate (Asprey et al 2001).

Un criterio per la discriminazione di modelli, è stato proposto da Hunter Reiner nel 1965 (Chen et. al 2003), ed è basato sull'ottimizzazione dinamica di una funzione obiettivo Φ_T che definisce

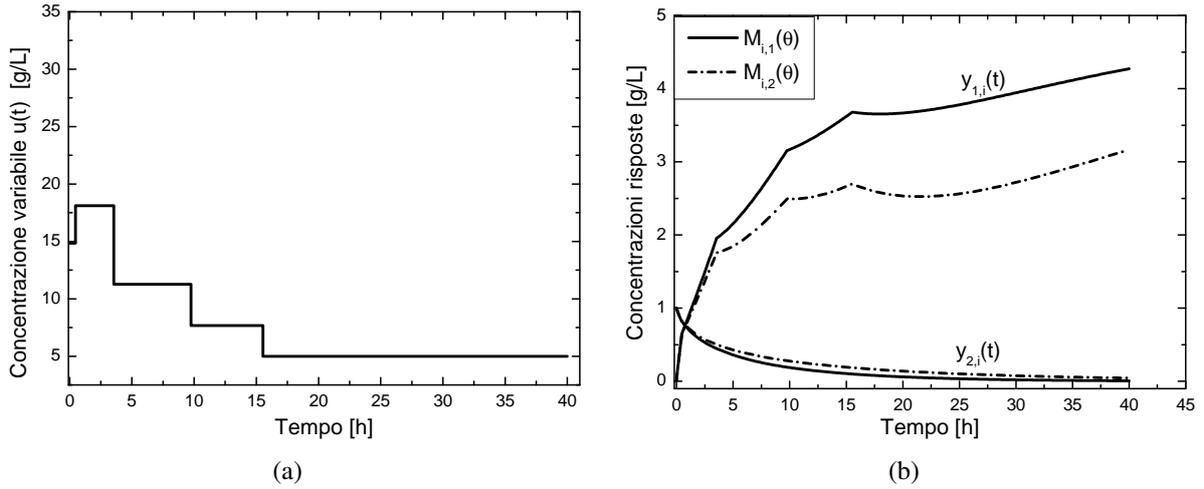


Figura 2.5: Esempio di MBDoe per la discriminazione di due modelli proposti. Nella figura (a) si nota il profilo della variabile controllata in modo tale da massimizzare la distanza di distinguibilità (b) tra i due modelli $\mathbf{M}_1(\boldsymbol{\theta})$ e $\mathbf{M}_2(\boldsymbol{\theta})$.

il seguente vettore degli esperimenti (2.29):

$$\boldsymbol{\varphi} = \arg \max_{\boldsymbol{\varphi} \in \Phi_T} \int_{t_f}^{t_0} \sum_{l=1}^{N_m} \sum_{l'=l+1}^{N_m} \left[\hat{\mathbf{y}}_l(\boldsymbol{\varphi}, \hat{\boldsymbol{\theta}}, t) + \right. \\ \left. - \hat{\mathbf{y}}_{l'}(\boldsymbol{\varphi}, \hat{\boldsymbol{\theta}}', t) \right]^T \mathbf{W} \left[\hat{\mathbf{y}}_l(\boldsymbol{\varphi}, \hat{\boldsymbol{\theta}}, t) - \hat{\mathbf{y}}_{l'}(\boldsymbol{\varphi}, \hat{\boldsymbol{\theta}}', t) \right] dt \quad (2.29)$$

dove $\hat{\mathbf{y}}_i$ sono le risposte dei modelli, $\hat{\boldsymbol{\theta}}_i$ il set dei parametri, N_m il numero di modelli proposti, \mathbf{W} la matrice dei pesi. L'ottimizzazione di tale funzione definisce le condizioni sperimentali per ottenere in uscita il vettore $\boldsymbol{\varphi}$ in modo tale da massimizzare la distanza ε_{Φ_D} di distinguibilità.

A titolo esemplificativo si possono osservare in Figura 2.5 i profili delle variabili controllate $\mathbf{u}(t)$, dopo il design del vettore $\boldsymbol{\varphi}$ tramite l'ottimizzazione di (2.29), e i profili delle risposte predette dai due modelli alle condizioni di design. Si noti come il design delle condizioni di $\mathbf{u}(t)$ massimizzi la distanza di distinguibilità tra i profili dei due modelli proposti $\mathbf{M}_1(\boldsymbol{\theta})$ e $\mathbf{M}_2(\boldsymbol{\theta})$. Esistono ulteriori criteri e funzioni obiettivo per il MBDoe al fine di discriminare diversi modelli, oltre al criterio Hunter-Reiner ad esempio si potrà utilizzare il criterio Buzzi-Ferraris e Forzatti (1984) il quale è stato sviluppato per calcolare gli ingressi dinamici $\mathbf{u}(t)$ tramite una riformulazione di un problema di progettazione degli esperimenti come se fosse un problema di controllo ottimale.

Capitolo 3

Analisi preliminare dei modelli proposti per la malattia di von Willebrand

In questo capitolo verranno analizzate le problematiche legate alla struttura e all'identificabilità parametrica dei modelli fisiologici che descrivono la VWD. Dopo una descrizione dei dati disponibili verranno presentate le analisi preliminari di sensibilità ed informazione dei modelli proposti secondo le tecniche formulate nel capitolo precedente.

3.1 Analisi dei dati disponibili

In riferimento alle diverse modalità di misura del VWF (§1.3), i dati a disposizione, forniti dal Dipartimento di Scienze Cardiologiche, Toraciche e Vascolari dell'Università di Padova, sono rappresentati dalle misure di antigene (VWF:AG) e di binding al collagene (VWF:CB) le quali possono essere divise in due categorie:

- profili medi per i pool di dati di soggetti normali 0, soggetti normali non-0, pazienti affetti da VWD tipo 2B, pazienti affetti da VWD tipo Vicenza;
- valori individuali per soggetti normali 0 (24 soggetti), soggetti non-0 (18 soggetti), pazienti VWD tipo 2A (1 soggetto), pazienti VWD tipo 2B (7 soggetti) e pazienti VWD tipo Vicenza (10 soggetti).

Le misure di VWF:AG rappresentano la concentrazione totale del VWF indipendentemente dalla grandezza delle unità multimeriche, mentre il VWF:CB indica la presenza dei multimeri ad elevato peso molecolare nel plasma, molecole che hanno maggiore affinità alle piastrine. Gli esperimenti vengono condotti tramite un'iniezione sottocutanea di 3 $\mu\text{g}/\text{kg}$ di DDAVP; i punti sperimentali di concentrazioni in funzione del tempo vengono riportati tramite il prelievo in diversi istanti prestabiliti. Dopo la somministrazione della DDAVP avvengono solitamente tre meccanismi:

- rilascio della macromolecola del VWF;
- proteolisi dei multimeri ad alto peso molecolare in specie di dimensioni inferiori, ad opera dell'enzima ADAMTS13; questa fase è influenzata dalla tipologia individuale del gruppo

sanguigno (AB0);

- clearance, o eliminazione dei multimeri dal plasma sanguigno, che solitamente avviene a livello del fegato.

I dati medi caratteristici per VWF:AG e VWF:CB sono riportati in Figura 3.1. I soggetti sani 0 e non-0 presentano una distribuzione normale dei multimeri di VWF, dove sia le specie ad alto peso molecolare che quelle a basso peso sono presenti nel plasma, infatti il rapporto tra VWF:AG e VWF:CB è molto vicino a 1 (Figura 3.2). Tuttavia i soggetti non-0 presentano un incremento di concentrazione del VWF maggiore del 25% ed una minore velocità di proteolisi. Questo aspetto innalza significativamente il rischio di complicazioni vascolari e malattie come trombosi e ischemie cardiache. Si ricordi il significato delle misure di VWF:AG, che rappresentano la concentrazione di tutti i multimeri del VWF, e VWF:CB le quali descrivono le concentrazioni dei soli multimeri ad alto peso molecolare (§1.5).

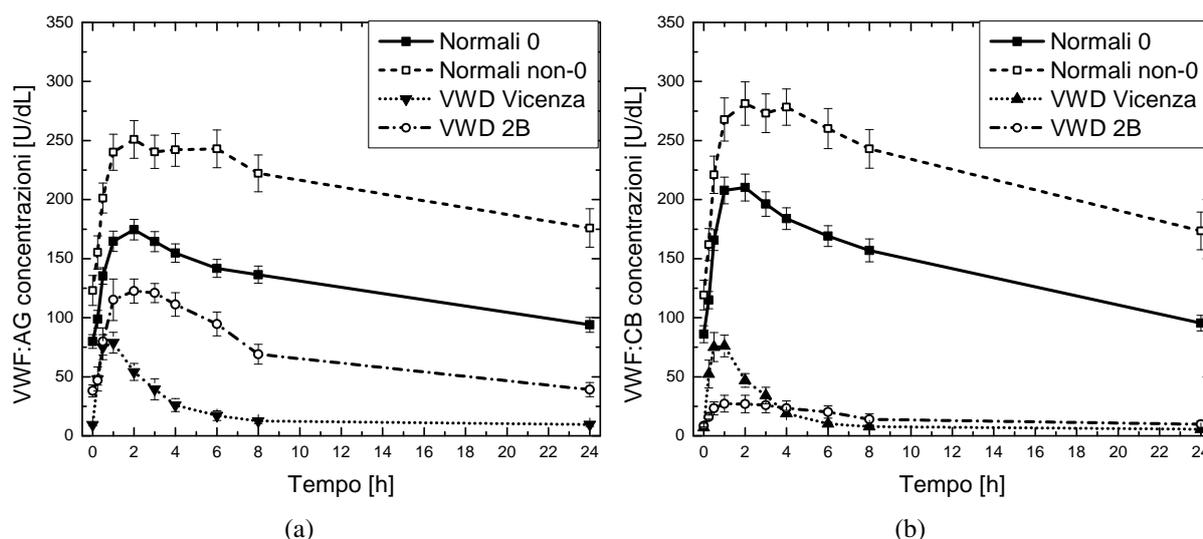


Figura 3.1: Dati medi di VWF:AG (a) e VWF:CB (b) per i soggetti normali 0 e non-0 e per i pazienti VWD di tipo Vicenza e 2B. Le bande rappresentano la deviazione standard del pool sui dati medi; le linee hanno solamente lo scopo di congiungere i punti sperimentali ed aiutare la rappresentazione grafica.

La VWD di tipo Vicenza è caratterizzata da una maggiore e più alta velocità di clearance (Figura 3.1a e b), a questo fatto conseguono concentrazioni di VWF più basse nel plasma. Nella tipologia 2B (Figura 3.1b) invece, la totale mancanza nei pazienti dei multimeri ad alto peso molecolare è causata da una maggiore e spontanea affinità di tali multimeri alle piastrine, anche se essi vengono prodotti normalmente.

Per quanto riguarda la tipologia 2A si manifesta un percorso differente per la fase di proteolisi dei multimeri. In questi pazienti la maggiore attività proteolitica dell'ADAMTS13 caratterizza

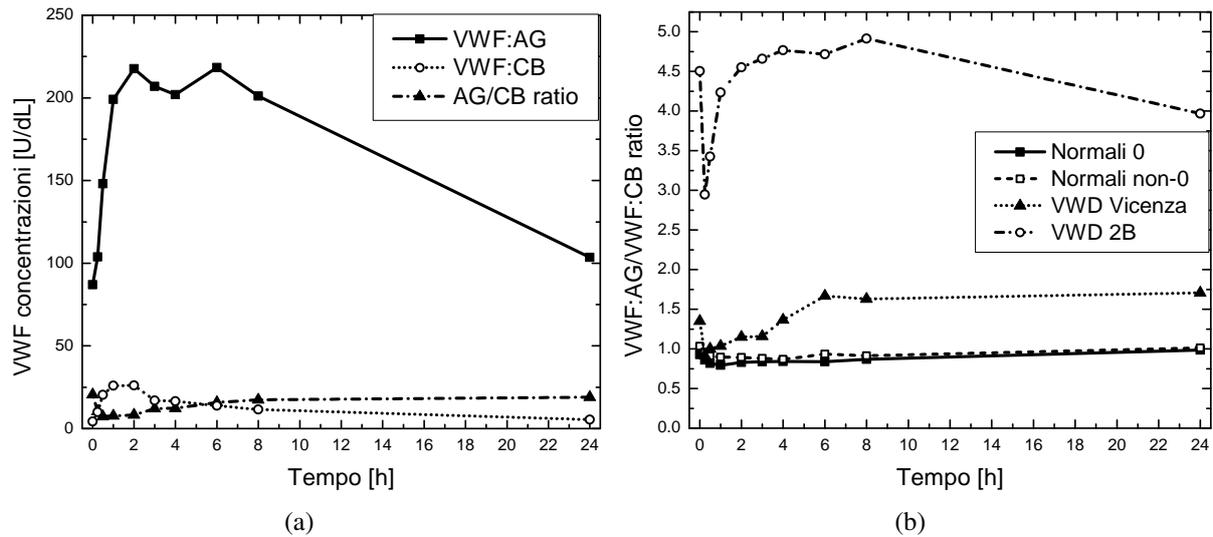


Figura 3.2: Valori di antigene, collagene e rispettivo rapporto per il paziente 2A (a); rapporti dei dati medi di VWF:AG/VWF:CB per i soggetti normali e affetti da VWD (b).

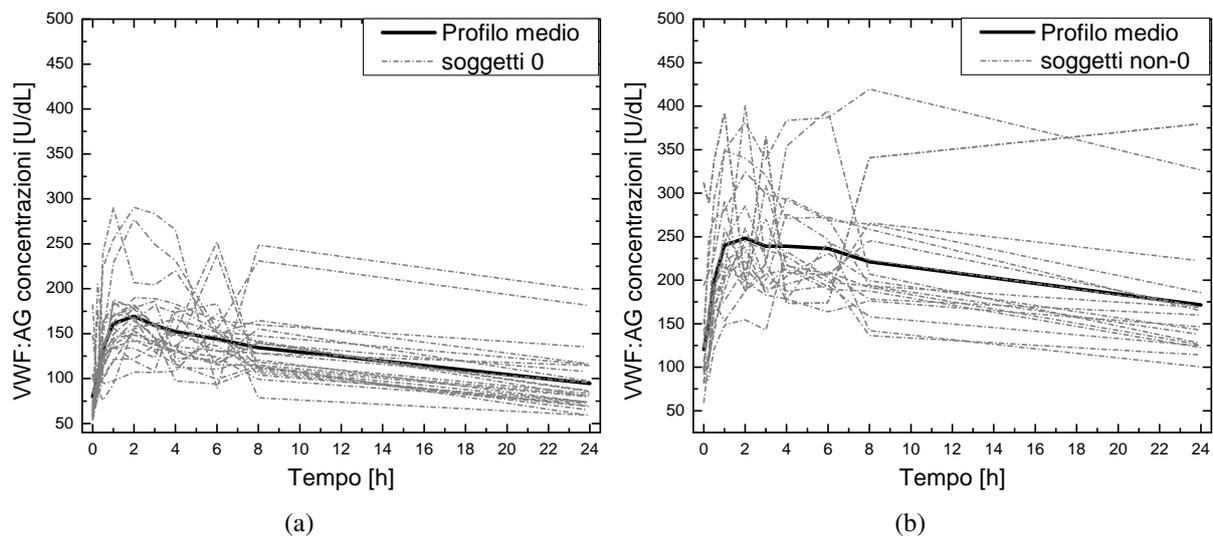


Figura 3.3: Valori di antigene disponibili per tutti i soggetti normali 0 (a) e non-0 (b). La linea continua rappresenta la congiunzione dei punti sperimentali medi per ogni caso.

una progressiva degradazione dei multimeri ad alto peso molecolare (Figura 3.1b). Un aspetto interessante da osservare è che il rapporto tra VWF:AG e VWF:CB (Figura 3.2a) nel tipo 2A è molto alto rispetto ai soggetti normali, ed ai pazienti affetti da VWD 2B e Vicenza (Figura 3.2b), nelle quali sono presenti basse concentrazioni di multimeri VWF ad alto peso molecolare.

I profili del VWF:AG presentano sia per i soggetti 0 (Figura 3.3a) che per i soggetti non-0 (Figura 3.3b) un'alta variabilità intra-individuale. Si notino inoltre le forti oscillazioni delle concentrazioni.

Un altro aspetto di notevole importanza risiede nella diversità nei valori delle concentrazioni

basali medie sia per i soggetti normali che per i pazienti affetti da VWD. In Tabella 3.1 vengono riportati i valori basali di VWF:AG e VWF:CB per ogni tipologia e la varianza associata alle misure.

Tabella 3.1: Valori basali di antigene e collagene in [U/dL] di VWF per i diversi casi di pazienti normali 0 e non-0 e affetti da VWD, con le rispettive deviazioni standard sulle misure.

	0	Non-0	VWD Vicenza	VWD 2A*	VWD 2B
VWF:AG	79.84 ±5.72	123.07 ±12.72	9.43 ±0.90	87.0 ±5.13	38.06 ±5.13
VWF:CB	86.13 ±7.20	119.14 ±12.60	6.98 ±0.83	4.25 ±2.14	8.45 ±2.14

* I valori per la tipologia VWD 2A sono dati provenienti da un solo paziente e quindi non sono da considerarsi valori medi.

3.2 Formulazione della struttura dei modelli candidati

Studi precedenti (Gallinaro *et al.*, 2008, Galvanin, 2012a) sono stati condotti su dei semplici modelli farmacocinetici con lo scopo di analizzare la loro adeguatezza nella descrizione del comportamento dei profili medi di VWF:AG e VWF:CB, e per poter effettuare delle analisi preliminari di sensitività investigando l'effetto principale dei parametri del modello sulle risposte misurate. Il limite di tali modelli è quello di non rappresentare in modo opportuno i complessi meccanismi biologici della fase di proteolisi (Casonato *et al.*, 2002), ma di limitarsi alle sole fasi di secrezione ed eliminazione del VWF. Nella formulazione strutturale dei nuovi modelli sono stati considerati i seguenti percorsi (Galvanin, 2012b):

- il rilascio delle macromolecole multimeriche, multimeri *Super Ultra Large* (SUL): è supposto che questi multimeri non vengano rilasciati direttamente nel flusso sanguigno e quindi non possono essere rilevati con delle misure;
- la proteolisi dei SUL in multimeri ad alto peso molecolare, multimeri *Ultra Large* (UL), multimeri in medio peso molecolare *High molecular weight* (H), e in multimeri a basso peso molecolare *Low molecular weight* (L);
- la riduzione dei multimeri UL in H, la riduzione dei multimeri H in L e la riduzione dei UL in L;
- l'eliminazione dei UL, H ed L, la cui costante di velocità è mantenuta indipendente rispetto alle dimensioni dei multimeri.

Dopo diverse rielaborazioni ed analisi investigative si è arrivati alla formulazione delle strutture di due modelli candidati che nel lavoro di Galvanin (2012c) vengono chiamati modelli 6M

e 7. La descrizione numerica del set di equazioni, delle strutture e significato fisiologico dei parametri dei modelli vengono riportate in dettaglio nel paragrafo successivo.

3.3 Modelli proposti

I due modelli proposti, 6M e 7, tra di loro presentano l'unica differenza nella descrizione del numero di classi multimeriche: nel modello 7 vengono accorpati i canali dei multimeri UL ed H in un unico canale proteolitico con il solo scopo di ridurre il numero dei parametri del modello. Poiché con il modello 6M si ha una rappresentazione dettagliata di ogni distinta classe multimerica, la trattazione seguente viene eseguita per entrambi i modelli.

Le strutture dei modelli sono rappresentate in Figura 3.4a per il modello 6M e in Figura 3.4b per il modello 7. Il significato fisiologico dei parametri e dei flussi associati viene descritto in Tabella 3.2 per il modello 6M ed in Tabella 3.3 per il modello 7.

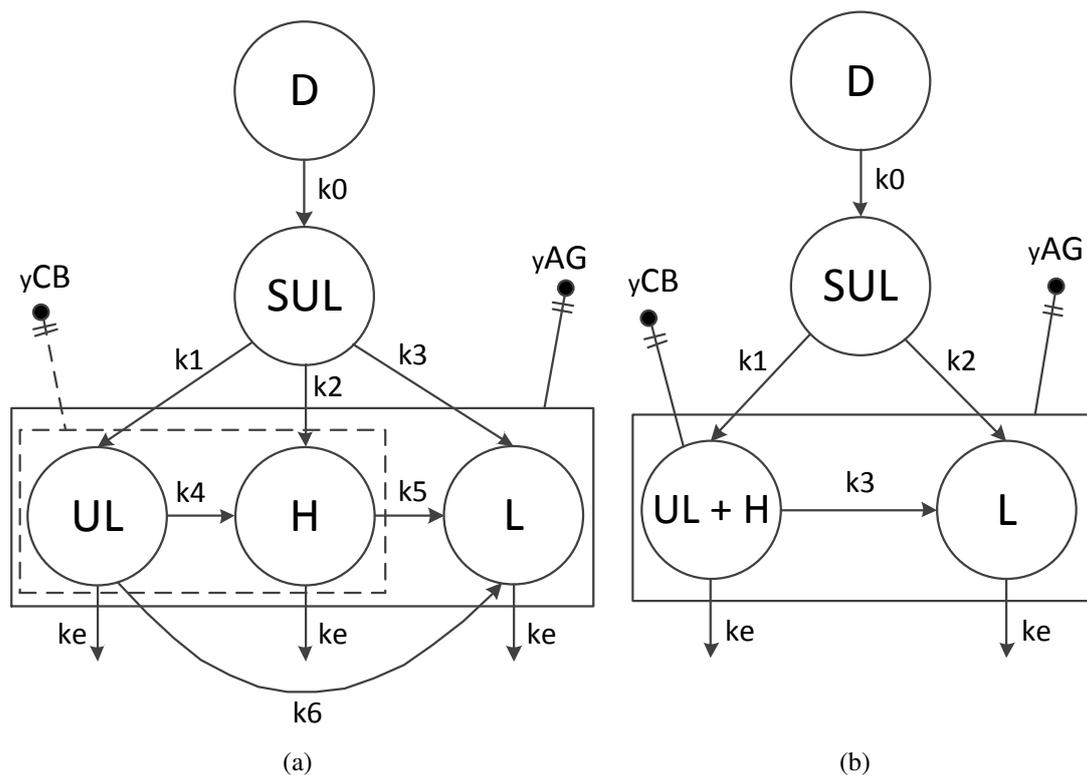


Figura 3.4: Strutture dei modelli proposti in termini di parametri coinvolti nei percorsi fisiologici: (a) modello 6M, (b) modello 7. Ogni blocco a forma di cerchio rappresenta rispettivamente una fase del meccanismo: D è l'esperimento dato dall'iniezione di DDAVP, SUL il rilascio conseguente della macromolecola glicoproteica di VWF, UL i multimeri ad alto peso molecolare, H i multimeri a medio peso molecolare ed L i multimeri a basso peso molecolare. I riquadri rappresentano l'accessibilità delle singole classi multimeriche tramite le misure disponibili di antigene y^{AG} e di collagene y^{CB} .

Tabella 3.2: Descrizione dei flussi di massa coinvolti nella formulazione del modello 6M.

Fasi	Flussi	Parametri	Descrizione
Rilascio	F_0	k_0, D, t_{max}	Rilascio di multimeri SUL
Proteolisi	F_1	k_1	Riduzione di multimeri SUL a UL
	F_2	k_2	Riduzione di multimeri SUL a H
	F_3	k_3	Riduzione di multimeri SUL a L
	F_4	k_4	Riduzione di multimeri UL a H
	F_5	k_5	Riduzione di multimeri H a L
	F_6	k_6	Riduzione di multimeri UL a L
Clearance	F_7	k_e	Eliminazione di multimeri UL
	F_8	k_e	Eliminazione di multimeri H
	F_9	k_e	Eliminazione di multimeri L

Il set di equazioni rappresentanti il modello 6M sono:

$$\begin{aligned} \frac{dx^{SUL}}{dt} &= k_0 D \exp^{-k_0(t-t_{max})} - k_1 (x^{SUL} - x_b^{SUL}) - k_2 (x^{SUL} - x_b^{SUL}) - k_3 (x^{SUL} - x_b^{SUL}) \\ &= F_0 - F_1 - F_2 - F_3 \end{aligned} \quad (3.1)$$

$$\begin{aligned} \frac{dx^{UL}}{dt} &= k_1 (x^{SUL} - x_b^{SUL}) - k_4 (x^{UL} - x_b^{UL}) - k_e (x^{UL} - x_b^{UL}) - k_6 (x^{UL} - x_b^{UL}) \\ &= F_1 - F_4 - F_7 - F_6 \end{aligned} \quad (3.2)$$

$$\begin{aligned} \frac{dx^H}{dt} &= k_2 (x^{SUL} - x_b^{SUL}) + k_4 (x^{UL} - x_b^{UL}) - k_5 (x^H - x_b^H) - k_e (x^H - x_b^H) \\ &= F_2 - F_4 - F_5 - F_8 \end{aligned} \quad (3.3)$$

$$\begin{aligned} \frac{dx^L}{dt} &= k_3 (x^{SUL} - x_b^{SUL}) + k_5 (x^H - x_b^H) - k_e (x^L - x_b^L) + k_6 (x^{UL} - x_b^{UL}) \\ &= F_3 + F_5 + F_6 - F_9 \end{aligned} \quad (3.4)$$

dove x è rispettivamente il numero di unità di SUL (*Super Ultra Large multimers*), UL (*Ultra Large multimers*), H (*High multimers*) e L (*Low multimers*) in unità [U] presenti nel plasma.

Le condizioni basali iniziali per $t = 0$, indicate di seguito dal pedice b , vengono definite dal vettore:

$$\mathbf{x}^0 = \mathbf{x}(0) = [x_b^{SUL} \quad x_b^{UL} \quad x_b^H \quad x_b^L]^T = [0 \quad 0 \quad x_b^H \quad y_b^{AG} V_d - x_b^H]^T \quad (3.5)$$

dove V_d è il volume di distribuzione posto uguale a 4 L (Menache *et al.*, 1996), mentre la durata dell'esperimento τ è sempre fissata a 24 h. Le concentrazioni osservate y^{AG} (3.6), concentra-

zione di VWF antigene, e y^{CB} (3.7), concentrazione di VWF collagene, quest'ultimo incluso nell'equazione di correzione la quale ha solamente lo scopo di correggere la risposta predetta y^{CB} in concentrazione reale osservabile $y^{CB'}$ (3.8):

$$y^{AG} = (x^{UL} + x^H + x^L) / V_d \quad (3.6)$$

$$y^{CB} = (x^{UL} + x^H) / V_d \quad (3.7)$$

$$y^{CB'} = ky^{CB} \frac{y_b^{AG}}{y_b^{CB}} = ky^{CB} \left(1 + \frac{x_b^L}{x_b^H} \right). \quad (3.8)$$

Il set di parametri da stimare è: $\theta = [k_0, k_1, k_2, k_3, k_4, k_5, k_6, k_e, D, t_{max}]$ con i due parametri aggiuntivi della correzione k e y_b^{CB} .

Per quanto riguarda il modello 7 il set di equazioni diventa (3.9-3.15):

$$\begin{aligned} \frac{dx^{SUL}}{dt} &= k_0 D \exp^{-k_0(t-t_{max})} - k_1 (x^{SUL} - x_b^{SUL}) - k_2 (x^{SUL} - x_b^{SUL}) \\ &= F_0 - F_1 - F_2 \end{aligned} \quad (3.9)$$

$$\begin{aligned} \frac{dx^{UL+H}}{dt} &= k_1 (x^{SUL} - x_b^{SUL}) - k_3 (x^{UL+H} - x_b^{UL+H}) - k_e (x^{UL+H} - x_b^{UL+H}) \\ &= F_1 - F_3 - F_4 \end{aligned} \quad (3.10)$$

$$\begin{aligned} \frac{dx^L}{dt} &= k_2 (x^{SUL} - x_b^{SUL}) + k_3 (x^H - x_b^H) - k_e (x^L - x_b^L) \\ &= F_2 + F_3 - F_5 \end{aligned} \quad (3.11)$$

dove le variabili coinvolte sono le stesse del modello precedente, fatta eccezione solamente per il raggruppamento in un solo comparto dei multimeri UL+L.

Le condizioni basali iniziali diventano:

$$\mathbf{x}^0 = \mathbf{x}(0) = [x_b^{SUL}, x_b^{UL+H}, x_b^L]^T = [0, x_b^{UL+H}, y_b^{AG} V_d - x_b^{UL+H}]^T \quad (3.12)$$

dove V_d è il volume di distribuzione posto uguale a 4 L (Menache *et al.*, 1996), τ la durata dell'esperimento pari a 24 h. Le concentrazioni osservate y^{AG} (3.13), y^{CB} (3.14) e l'equazioni di correzione (3.15):

$$y^{AG} = (x^{UL+H} + x^L) / V_d \quad (3.13)$$

Tabella 3.3: Descrizione dei flussi di massa coinvolti nella formulazione del modello 7.

Fasi	Flussi	Parametri	Descrizione
Rilascio	F_0	k_0, D, t_{max}	Rilascio di multimeri SUL
Proteolisi	F_1	k_1	Riduzione di multimeri SUL a UL+H
	F_2	k_2	Riduzione di multimeri SUL a L
	F_3	k_3	Riduzione di multimeri UL+H a L
Clearance	F_4	k_e	Eliminazione di multimeri UL+H
	F_5	k_e	Eliminazione di multimeri L

$$y^{CB} = (x^{UL+H}) / V_d \quad (3.14)$$

$$y^{CB'} = k_{y^{CB}} \frac{y_b^{AG}}{y_b^{CB}} = k_{y^{CB}} \left(1 + \frac{x_b^L}{x_b^{UL+H}} \right). \quad (3.15)$$

Il set di parametri da stimare è: $\theta = [k_0, k_1, k_2, k_3, k_e, D, t_{max}]$ con i due parametri aggiuntivi di correzione k e y_b^{CB} .

Si può osservare in Figura 3.4 come le misure di VWF:AG corrispondono nei modelli alla somma delle concentrazioni delle specie multimeriche UL, H e L, mentre le misure di VWFG:CB corrispondono solamente alla somma delle concentrazioni di multimeri UL e H.

3.4 Stime preliminari dei modelli 6M e 7

Il software utilizzato per le stime e tutte le simulazioni successive in questo lavoro di tesi è stato gPROMS[®] ver3.2.1. Dopo una stima preliminare dei parametri, le stime successive sono state condotte normalizzando i parametri del modello per facilitare la soluzione numerica durante le operazioni di stima (*Parameter Estimation*), poiché i parametri stessi possono assumere valori di ordini di grandezza diversi. La normalizzazione dei parametri consiste nel dividere ciascun elemento di θ per il fattore di normalizzazione μ_i riferito a ciascun parametro:

$$\theta_i = \frac{\hat{\theta}_i}{\mu_i} \quad i = \dots N_\theta \quad (3.16)$$

si considera poi il nuovo set di parametri $\Theta = [\theta_1 \dots \theta_{N_\theta}]$. I fattori di normalizzazione possono essere i valori nominali adottati nella stessa attività di design, in questo caso si è utilizzato inizialmente il set θ_i^0 effettuato da una stima preliminare in Galvanin (2012b).

3.4.1 Stime effettuate per i soggetti normali 0 e non-0

In Tabella 3.4 vengono riportate le stime dei parametri, i valori normalizzati e i limiti superiori ed inferiori degli intervalli di stima del modello 6M sia per i soggetti 0 che non-0. Una volta ottenuta la prima stima si è compiuto un percorso iterativo di affinamento dei parametri, utilizzando come valori iniziali della stima successiva i valori normalizzati della stima precedente. Il modello di varianza utilizzato nelle stime è stato di tipo costante pari a $\omega = 3$ U/dL.

Con le stime effettuate si ottengono le risposte di VWF:AG e VWF:CB di Figura 3.5a e b, e le concentrazioni predette dei multimeri di Figura 3.5c e d. Si osserva come le curve di antigene e collagene rispecchino i dati sperimentali: in entrambi i casi a curva VWF:CB è al di sopra della VWF:AG; il picco di concentrazione è in corrispondenza delle 2 h, ed in entrambi i casi il valore a fine esperimento non ritorna alle stesse concentrazioni per $t = 0$. Questo significa che la clearance è molto lenta ed è necessario un tempo maggiore delle 24 h per ritornare alle

Tabella 3.4: Stime dei parametri, guess iniziali e intervalli di stima per il modello 6M.

SOGGETTI 0					
Stime	Valori normalizzati	Guess	Low bound	Up bound	
k_0	0.0574	0.99	1	0	5
k_1	0.0082	1.005	1	0	5
k_2	0.0451	0.99	1	0	5
k_3	0.004	1.005	1	0	5
k_4	0.0012	0.993	1	0	5
k_5	8.41×10^{-5}	1.00	1	0	5
k_6	0.0238	0.99	1	0	5
k_e	0.0014	1.065	1	0	5
D	60.74	1.00	1	0	5
t_{\max}	33.70	1.00	1	0	5
k	1.04	1.00	1	0.7	1.1
y_b^{CB}	48.82	0.99	1	0.5	1.5
SOGGETTI NON-0					
Stime	Valori normalizzati	Guess	Low bound	Up bound	
k_0	0.0413	0.999	1	0	5
k_1	6.73×10^{-9}	0.006	1	0	5
k_2	0.0836	1.038	1	0	5
k_3	0.0288	0.989	1	0	5
k_4	0.0014	1.46	1	0	5
k_5	1.46×10^{-4}	1.04	1	0	5
k_6	0.0719	3.21	1	0	5
k_e	6.61×10^{-4}	1.001	1	0	5
D	118.39	0.999	1	0	5
t_{\max}	37.10	1.001	1	0	5
k	0.958	0.998	1	0.8	1.1
y_b^{CB}	66.54	0.993	1	0.5	1.5

Tabella 3.5: Stime dei parametri, guess iniziali e intervalli di stima per il modello 7.

SOGGETTI 0					
Stime	Valori normalizzati	Guess	Low bound	Up bound	
\mathbf{k}_0	0.0607	1.057	1.000	0	5
\mathbf{k}_1	0.0527	1.033	1.060	0	5
\mathbf{k}_2	0.0080	1.998	2.168	0	5
\mathbf{k}_3	2.88×10^{-4}	0.240	0.209	0	5
\mathbf{k}_e	1.426×10^{-3}	1.019	1.061	0	5
\mathbf{D}	57.79	0.952	1.000	0	5
\mathbf{t}_{\max}	32.40	0.962	1.000	0	5
\mathbf{k}	1.0574	1.017	1.021	0.7	1.1
\mathbf{y}_b^{CB}	51.49	1.055	1.069	0.5	1.5
SOGGETTI NON-0					
Stime	Valori normalizzati	Guess	Low bound	Up bound	
\mathbf{k}_0	0.0413	0.999	1.000	0	5
\mathbf{k}_1	0.0893	1.085	1.081	0	5
\mathbf{k}_2	0.0231	0.794	0.796	0	5
\mathbf{k}_3	1.46×10^{-3}	0.146	0.145	0	5
\mathbf{k}_e	6.61×10^{-4}	1.001	1.003	0	5
\mathbf{D}	118.39	1.000	1.000	0	5
\mathbf{t}_{\max}	37.11	1.000	1.000	0	5
\mathbf{k}	0.9579	0.998	0.998	0.9	1.1
\mathbf{y}_b^{CB}	71.06	1.061	1.061	0.5	1.5

condizioni basali. Per quanto riguarda i multimeri dei pazienti non-0, i profili degli H ed L presentano valori di circa 100 U superiori rispetto ai pazienti 0 ed il numero dei UL è molto basso o quasi inesistente, aspetto dovuto al basso valore di k_1 (Tab. 3.4), inoltre nei soggetti 0 la curva sempre degli UL è presente con un picco più evidente, dell'ordine dei 20 U, e l'andamento degli H sembrerebbe avere una clearance propria leggermente più rapida. Purtroppo il modello è sovraparametrizzato e quindi dalla fase di stima non sono presenti le statistiche in uscita a causa di Φ^{ML} , poiché avendo tale funzione un andamento molto piatto, lo stimatore non è in grado di calcolare le matrici di varianza-covarianza e di correlazione dei parametri.

Nella Figura 3.6 sono rappresentate le entità dei diversi flussi predetti a fine esperimento, per le diverse fasi e per soggetti 0 (Figura 3.6a) e soggetti non-0 (Figura 3.6b), riferiti in *Area Under the Curve* (AUC):

$$AUC = \int_0^{\tau} F_i dt \quad (3.17)$$

dove τ è la durata dell'esperimento e F_i sono i flussi descritti in Tabella 3.2. Si nota come per i soggetti non-0 il flusso F1, corrispondente ai multimeri UL, sia molto basso e di conseguenza anche F4 (da UL a H), F6 (da UL a L) ed F7 (clearance UL).

Le stime condotte per il modello 7, nelle stesse modalità del modello 6M, sono riportate in

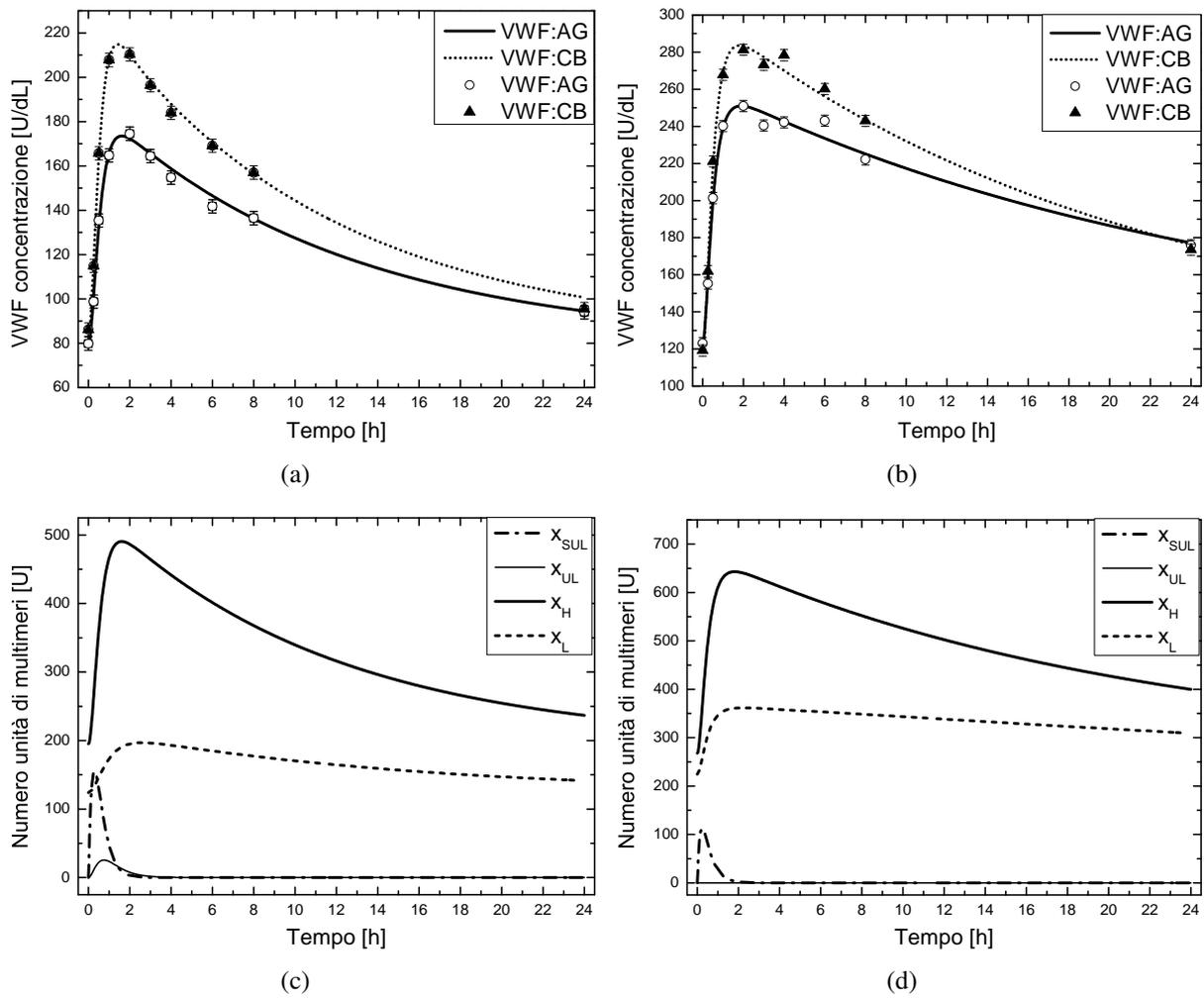


Figura 3.5: Risposte di antigene e di collagene predette dal modello 6M per i soggetti gruppo 0 (a) e gruppo non-0 (b); i punti rappresentano i corrispettivi set di dati sperimentali medi. Le figure inferiori rappresentano invece le unità multimeriche durante la durata dell'esperimento per pazienti 0 (c) e pazienti non-0 (d).

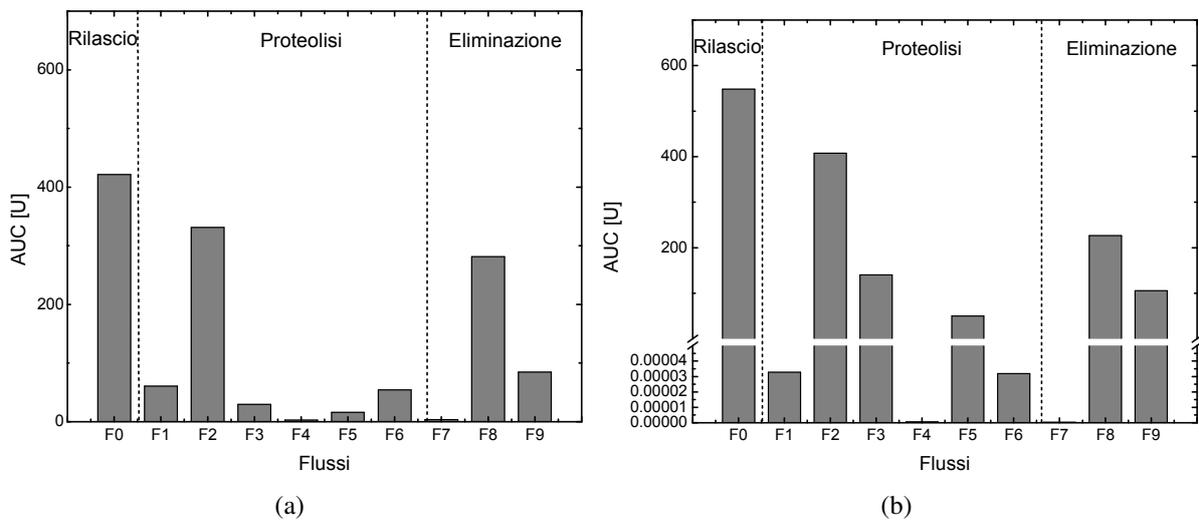


Figura 3.6: Flussi predetti dal modello 6M a fine esperimento, per i soggetti 0 (a) e non-0 (b). Il significato fisiologico di ciascun flusso è descritto in Tabella 3.2

Tabella 3.5. L'unica differenza tra i due modelli, come accennato nel §3.3, consiste nel raggruppare nel modello 7 i flussi dei multimeri UL e H del modello 6M, corrispondenti alle misure di collagene, ottenendo un unico flusso UL+H. Questo perché come visto in precedenza il canale dei multimeri UL è di difficile caratterizzazione, soprattutto per i soggetti gruppo non-0.

Passando dal modello 6M al modello 7 si riduce il numero di parametri da 12 a 9 (si confrontino Figura 3.4) e ciò, da un punto di vista teorico, dovrebbe rendere il modello più robusto in fase di stima. L'aspetto negativo però è quello di eliminare la differenziazione delle due diverse classi di multimeri UL e H, perdendo l'informazione sui singoli andamenti.

I profili delle risposte di VWF:AG e CB predette dal modello 7 presentano andamenti esattamente identici alle curve predette dal modello 6M. La stessa cosa vale per i multimeri ma in questo caso il profilo dei multimeri UL sarà sommato al profilo dei multimeri H. Questo conferma la motivazione dell'implementazione del modello 7 che, in fase di stima, è unicamente legata alla riduzione del numero di parametri ed ottenere delle stime più robuste. Per sintesi si omette di riportare i grafici del modello 7.

3.4.2 Stime effettuate per i pazienti affetti da VWD

Inizialmente si sono presi in considerazione i dati medi per i pazienti VWD di tipo Vicenza e 2B. Le stime sono state condotte sempre nelle stesse modalità del paragrafo precedente utilizzando come parametri per la normalizzazione le stime dei soggetti sani 0.

I valori degli intervalli di stima rispetto al parametro k vengono scelti, da tipologia a tipologia, in funzione dei rapporti tra antigene e collagene specifici, poiché il valore e la variabilità di tale parametro dipendono fortemente dai valori di VWF:AG e VWF:CB, i quali sono diversi per ogni tipologia (Figura 3.2). Nelle Tabelle 3.6 (VWD Vicenza) e 3.7 (VWD 2B) sono riportati i parametri, i valori normalizzati, i valori iniziali e gli intervalli di stima sia per il modello 6M che per il modello 7.

I profili delle risposte predette dai modelli 6M e 7 sono identiche, quindi di seguito si riporteranno solamente i grafici riferiti al modello 6M e qualora ci siano delle discrepanze verranno descritte e specificate.

Per quanto riguardano le concentrazioni di antigene e collagene ed il numero di unità multimeriche per i pazienti affetti da VWD Vicenza, i profili sono rappresentati in Figura 3.7. Il primo aspetto lampante è che le curve VWF:AG e VWF:CB sono molto vicine tra di loro con una clearance molto rapida, che inizia già verso le 2 h e con un ritorno al valore basale entro le 8-9 h

Tabella 3.6: Stime dei parametri, guess iniziali e intervalli di stima per soggetti di tipo Vicenza.

Modello 6M					
	Stime	Valori normalizzati	Guess	Low bound	Up bound
k_0	0.0584	1.0181	1.010	0	20
k_1	0.0352	4.2868	4.300	0	20
k_2	0.154	3.4236	3.400	0	20
k_3	0	7.97×10^{-7}	0.000	0	20
k_4	0.0237	19.729	20.00	0	20
k_5	0.00136	16.188	16.00	0	20
k_6	0.0887	3.7283	3.700	0	20
k_e	0.00789	5.6363	5.600	0	20
D	59.7	0.98297	0.970	0	20
t_{\max}	32.2	0.95688	0.950	0	20
k	0.719	0.69182	0.690	0.5	1.1
y_b^{CB}	5.58	0.8	0.800	0.8	1.2
Modello 7					
	Stime	Valori normalizzati	Guess	Low bound	Up bound
k_0	0.05951	1.04	1.040	0	20
k_1	0.19272	3.78	3.780	0	20
k_2	0.01346	3.37	3.360	0	20
k_3	0.00172	1.43	1.430	0	20
k_e	0.00769	5.49	5.490	0	20
D	58.134	0.957	0.960	0	20
t_{\max}	31.792	0.943	0.940	0	20
k	0.69267	0.666	0.670	0.5	1.1
y_b^{CB}	5.5842	0.8	0.800	0.8	1.2

dall'inizio dell'esperimento. Questa è una caratteristica tipica della tipologia Vicenza derivante dalla indistinguibilità dei multimeri. Si osserva una concentrazione con un picco ad 80 U/dL, livello inferiore rispetto ai 200-250 U/dL dei soggetti sani, dovuto ai valori dei multimeri rispettivi (Figura 3.7b).

La Figura 3.8 invece riporta i profili di antigene e collagene e le unità multimeriche per la VWD di tipo 2B. In questo caso i valori delle concentrazioni (Figura 3.8a) sono anch'essi inferiori rispetto ai soggetti sani, ma il divario tra la curva di VWF:AG e la VWF:CB è molto più marcato, segno che i multimeri L sono presenti in numero maggiore (Figura 3.8b). La clearance presenta una dinamica leggermente più lenta rispetto ai Vicenza e si ha un ritorno alle condizioni basali dei valori di VWF:AG e CB leggermente prima delle 24 h. Osservando le curve dei multimeri, sempre in Figura 3.8b, si nota un notevole picco nel rilascio di unità delle macromolecole SUL con valori che sono quasi il doppio rispetto alla VDW Vicenza (Figura 3.7b) e anche per i soggetti sani (Figura 3.5c e d).

Nel caso della tipologia 2A per la mancanza di un numero di dati significativi a causa della

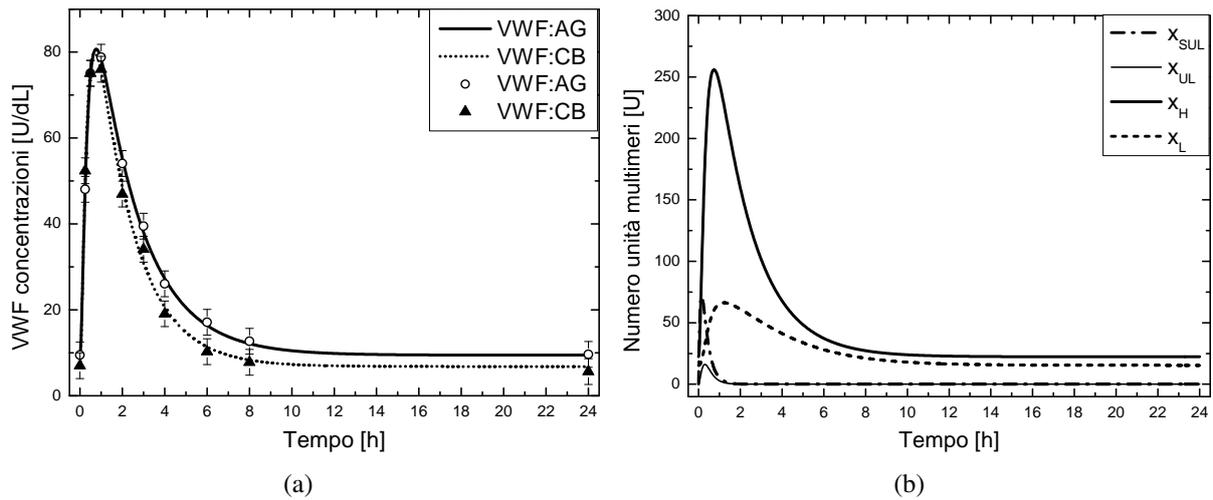


Figura 3.7: Risposte di antigene e collagene (a) e unità multimeriche (b) predette dal modello 6M per soggetti VWD di tipo 1.

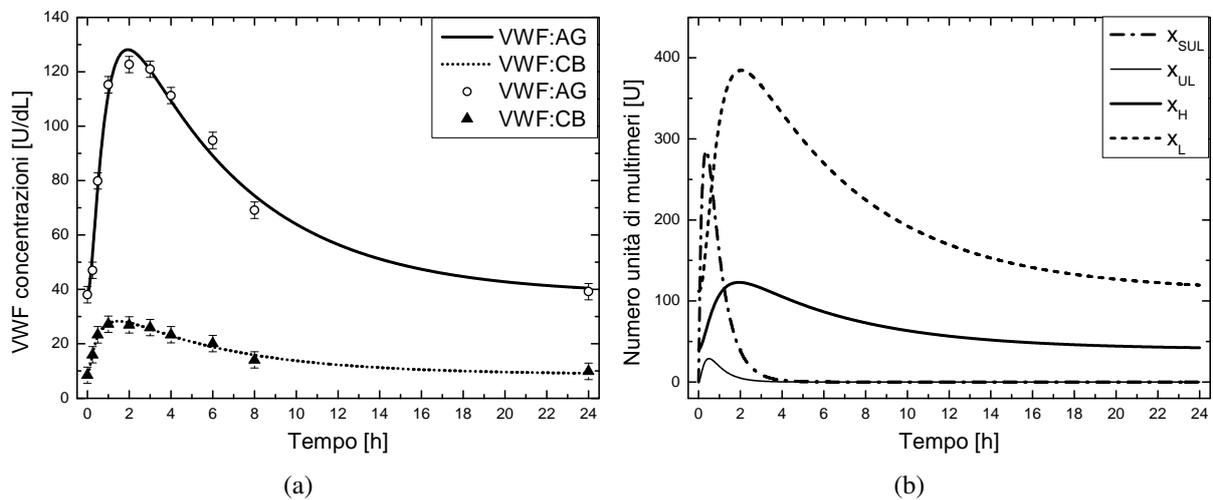


Figura 3.8: Risposte di antigene e collagene (a) e unità multimeriche (b) predette dal modello 6M per soggetti VWD di tipo 2B.

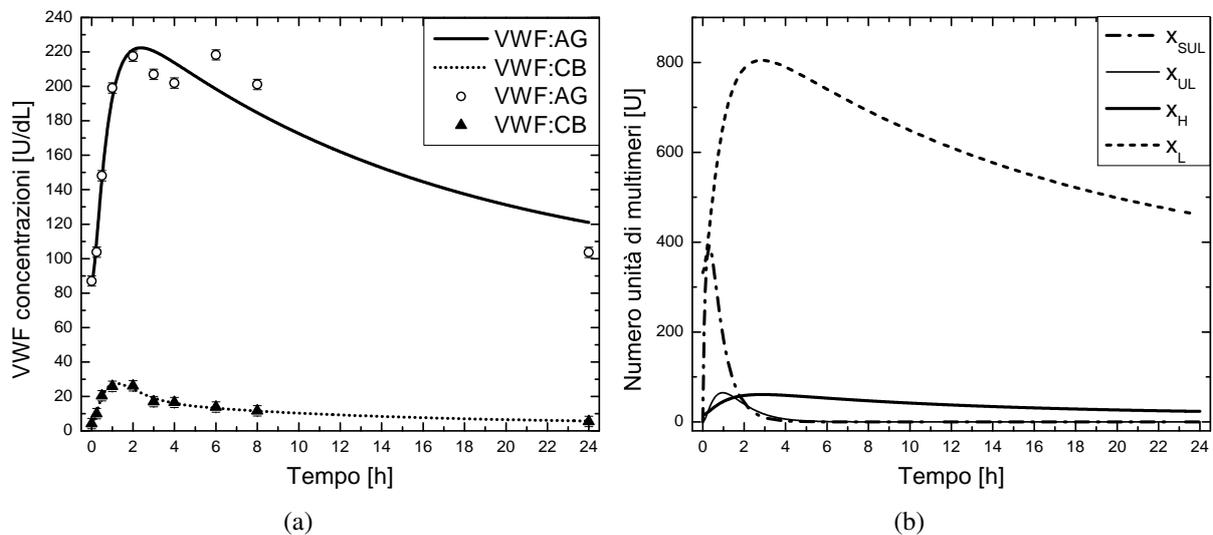


Figura 3.9: Risposte di antigene e collagene (a) e unità multimeriche (b) predette dal modello 6M per soggetto VWD di tipo 2A.

Tabella 3.7: Stime dei parametri, guess iniziali e limiti di convergenza per soggetti di tipo 2B.

Modello 6M					
	Stime	Valori normalizzati	Guess	Low bound	Up bound
k_0	0.0780	1.36	1.000	0	20
k_1	0.0167	2.04	2.340	0	20
k_2	0.00509	0.113	0.150	0	20
k_3	0.0023	0.575	2.760	0	20
k_4	0.00439	3.66	1.800	0	20
k_5	5.54×10^{-5}	0.659	6.040	0	20
k_6	0.1502	6.31	8.690	0	20
k_e	0.00284	2.02	1.500	0	20
D	44.98	0.741	1.000	0	20
t_{\max}	30.39	0.902	1.000	0	20
k	0.229	0.221	0.220	0.2	0.5
y_b^{CB}	10.05	1.19	0.880	0.8	1.2
Modello 7					
	Stime	Valori normalizzati	Guess	Low bound	Up bound
k_0	0.08144	1.42	1.000	0	20
k_1	0.00485	0.095	0.170	0	20
k_2	0.01918	4.79	7.150	0	20
k_3	0.00174	1.45	1.360	0	20
k_e	0.00281	2.01	1.570	0	20
D	41.749	0.687	1.000	0	20
t_{\max}	29.998	0.89	1.000	0	20
k	0.28515	0.274	0.270	0.2	0.5
y_b^{CB}	9.3889	1.11	1.200	0.8	1.2

rarietà della tipologia, non si hanno dei valori medi e quindi le stime effettuate valgono solo per il soggetto preso in considerazione. In Tabella 3.8 vengono riportate le stime parametriche per il modello 6M e 7, mentre in Figura 3.9 sono descritte le curve predette dal modello 6M di antigene e collagene (Figura 3.9a) e le unità multimeriche (Figura 3.9b). Per questa tipologia il divario tra le concentrazioni di VWF:AG e CB è molto più marcato nel caso della VWD 2B confermando il fatto che la produzione di multimeri ad elevato peso molecolare è ridotta, essendo tali multimeri centrali nel meccanismo emostatico (Casonato *et al.*, 1998). Questa caratteristica determina la disfunzione coagulativa caratteristica della tipologia 2A, infatti i multimeri H sono presenti in un numero molto basso e confrontabile con gli UL (Figura 3.9b). Un'altro aspetto importante è il valore delle concentrazioni di antigene comparabili ai soggetti normali 0, con uno stesso andamento normale dell'eliminazione che anche in questo caso a fine esperimento non riesce a ristabilire una concentrazione pari al livello basale. Come affermato in precedenza, avendo a disposizione dati riferiti ad un solo paziente di VWD 2A per la scarsità della tipologia, ai fini dello studio si assume tale paziente come soggetto medio.

Tabella 3.8: Stime dei parametri, guess iniziali e limiti di convergenza per soggetti di tipo 2A.

Modello 6M					
Stime	Valori normalizzati	Guess	Low bound	Up bound	
k_0	0.0975	1.698	1.000	0	20
k_1	0.006361	0.775	3.41	0	20
k_2	0.0014	0.0311	0.061	0	20
k_3	0.017187	4.296	11.34	0	20
k_4	0.00262	2.185	3.58	0	20
k_5	3.69×10^{-4}	4.393	16.78	0	20
k_6	0.01545	0.649	0.229	0	20
k_e	0.00109	0.785	0.306	0	20
D	35.923	0.591	1.00	0	20
t_{\max}	29.32	0.870	1.00	0	20
k	0.044	0.0426	0.04	0.03	0.3
y_b^{CB}	3.926	0.923	1.2	0.8	1.2
Modello 7					
Stime	Valori normalizzati	Guess	Low bound	Up bound	
k_0	0.0956	1.66	1.000	0	20
k_1	0.00283	0.055	0.409	0	20
k_2	0.02259	5.64	13.91	0	20
k_3	0.00502	4.18	2.58	0	20
k_e	0.0011	0.78	0.306	0	20
D	37.84	0.62	1.000	0	20
t_{\max}	29.32	0.87	1.000	0	20
k	0.0735	0.070	0.052	0.03	0.3
y_b^{CB}	3.424	0.80	1.200	0.8	1.2

3.5 Analisi di sensitività

A questo punto si effettua un'analisi di sensitività tramite il quale si potranno studiare gli effetti dei parametri sulle dinamiche delle risposte e dei diversi flussi dei modelli. Non essendoci differenza qualitativa tra i profili degli elementi della matrice di sensitività del modello 6M e quelle del modello 7, per sintesi verranno riportate solamente le curve rispetto al modello 6M. Infatti il modello 7 presenta dei flussi riferiti alla proteolisi che derivano dalla somma di due flussi del modello 6M, quindi gli andamenti delle curve corrispettive presentano un andamento che risulta essere la somma degli andamenti delle curve del modello 6M, questo aspetto si riflette anche nelle stime dei parametri, per esempio si osservino i parametri riferiti a k_1 e k_2 di Tabella 3.4 per il modello 6M con il valore di k_1 in Tabella 3.5 per il modello 7.

In riferimento alle figure di Figura 3.10 per i soggetti 0 e di Figura 3.11 per i soggetti non-0, si possono raggruppare i profili di sensitività seguendo la struttura delle diverse fasi del modello, e i corrispettivi parametri:

- alla fase di rilascio: k_0 , D e t_{max} ;
- alla fase di proteolisi: k_1 , k_2 , k_3 , k_4 , k_5 , k_6 ;
- alla fase eliminazione: k_e ;
- all'equazione di correzione: k e y_b^{CB} .

Per entrambi i soggetti normali 0 e non-0 e per VWF:AG e VWF:CB la fase del rilascio è dominata dal parametro k_0 che determina la secrezione dei multimeri SUL con una dinamica avente il picco massimo prima delle 2 h. Le costanti di tempo D e t_{max} hanno valori di sensitività molto bassi ma con un andamento simile rispetto a k_0 ed identico tra di loro.

La fase di proteolisi è caratterizzata dal parametro k_e , essendo la clearance la dinamica più lenta risulta essere quella dominante. I profili di sensitività di k_e rispecchiano l'andamento della velocità di eliminazione specifica di ogni soggetto: il picco per i soggetti 0 è tra le 10-12 h, quindi più veloce rispetto ai soggetti non-0 nei quali il punto massimo è spostato verso le 20-22 h.

I parametri della proteolisi presentano una sensitività diversa in base esse siano riferite all'antigene o al collagene. Per VWF:AG i parametri k_1 , k_2 e k_3 , sia per soggetti 0 che non-0, presentano un andamento identico: questo perché, osservando la struttura del modello in Figura 3.2, le misure di VWF:AG sono riferite alla somma dei multimeri UL, H ed L, quindi tale misura non fornisce alcuna informazione sui singoli corrispettivi parametri, ma sulla somma delle produzioni. Lo stesso vale per i parametri k_4 , k_5 e k_6 dove in questo caso è nulla, visto che sono dei parametri riferiti alla ripartizione tra i multimeri UL, H e L.

I profili rispetto ai parametri della proteolisi per il VWF:CB invece presentano andamenti diversi. Primo tra tutti si osservi l'andamento della curva rispetto a k_5 la quale è sovrapposta a quella di k_e : questo perché la proteolisi dei multimeri H a L può essere considerata alla stregua di un'eliminazione rispetto alle misure di VWF:CB, ricordando che il collagene rappresenta la somma dei multimeri UL e H. Le stesse considerazioni si possono compiere per k_6 , ma le quantità presenti dei multimeri UL (F1 di Figura 3.6) diminuiscono le sensitività rendendole nulle per i soggetti non-0 (Figura 3.11d). I parametri k_1 , k_2 e k_3 per il VWF:CB presentano dei picchi nelle prime 2 h, come i profili di sensitività riferiti al VWF:AG ma con andamenti diversi. Le misure di collagene danno delle informazioni sulla tipologia dei diversi multimeri rispetto all'antigene, cosa che invece non accade rispetto al parametro k_4 . Siccome tale parametro è riferito al flusso tra UL ed H nessuna misura disponibile fornisce informazione sul canale F4 (Tab. 3.2).

Il parametro della correzione k presenta sensitività nulla sia per le misure di VWF:AG che di VWF:CB; mentre y_b^{CB} ha sensitività nulla solo per VWF:AG, e poiché secondo l'equazione (3.8)

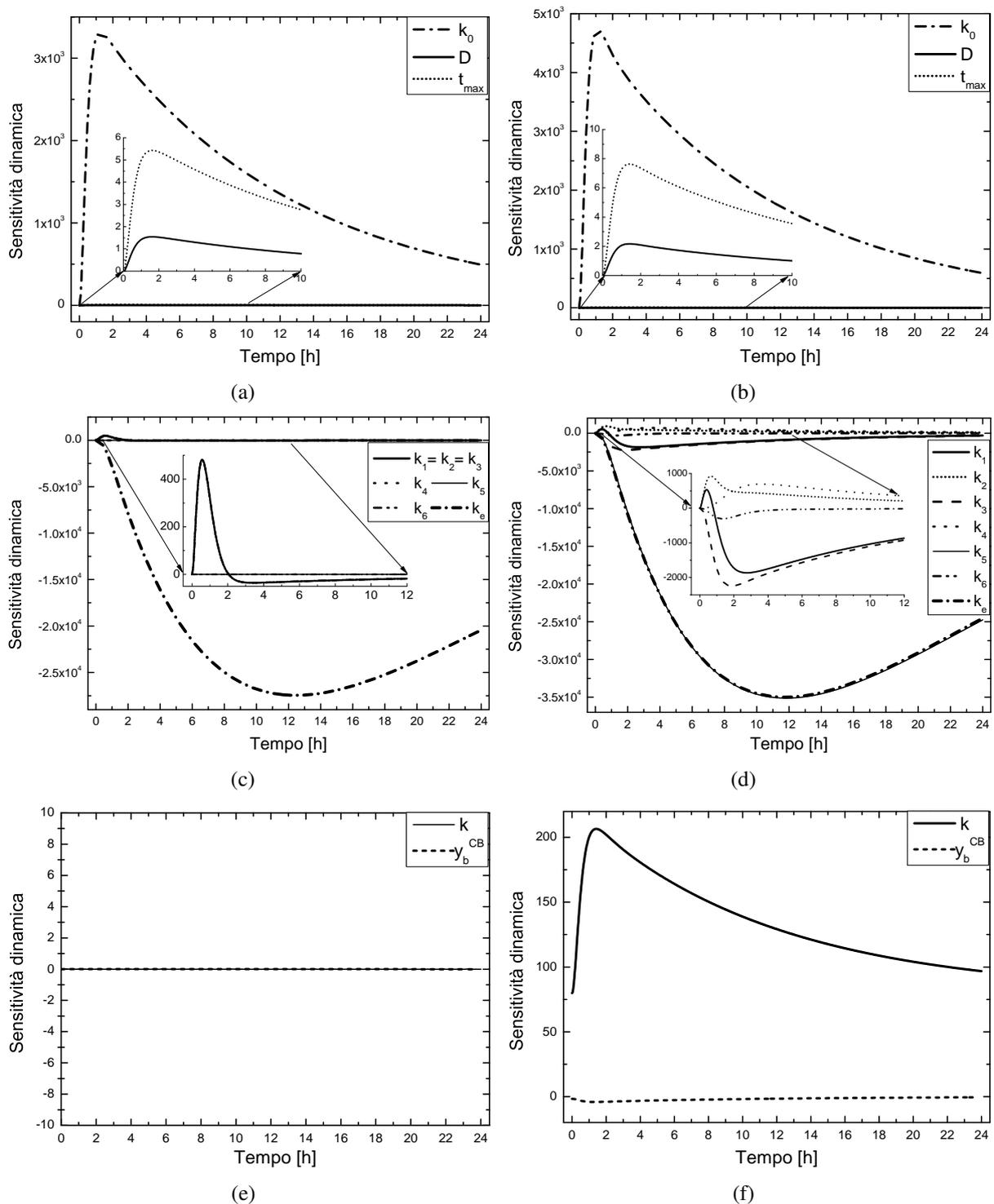


Figura 3.10: Curve di sensibilità per i soggetti gruppo 0; le figure a sinistra sono riferite all'antigene per le fasi di rilascio (a) proteolisi e clearance (c) e per i parametri di correzione (e), mentre quelle alla destra al collagene per le fasi di rilascio (b) proteolisi e clearance (d) e per i parametri di correzione (f).

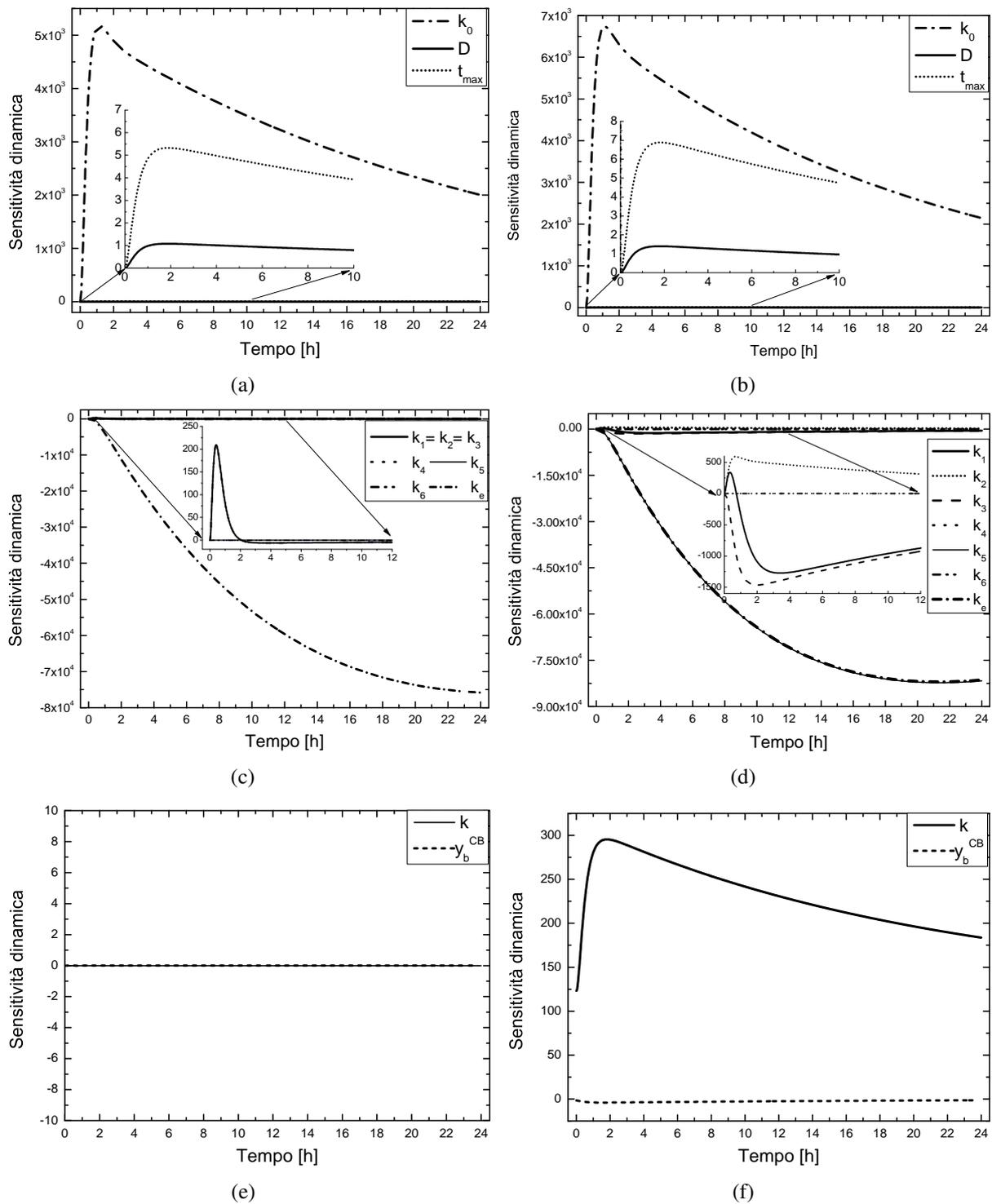


Figura 3.11: Curve di sensitività per i soggetti gruppo non-0; le figure a sinistra sono riferite all'antigene per le fasi di rilascio (a) proteolisi e clearance (c) e per i parametri di correzione (e), mentre quelle alla destra al collagene per le fasi di rilascio (b) proteolisi e clearance (d) e per i parametri di correzione (f).

esso dipende dal rapporto tra le unità multimeriche L e H, presenta un andamento non nullo per VWF:CB anche se di bassa intensità.

Effettuando la stessa analisi per i profili di sensitività riferite ai dati medi dei pazienti affetti da VWD, si trovano delle dinamiche con lo stesso andamento qualitativo dei profili. Quindi valgono le stesse considerazioni, con l'unica differenza che i profili seguono le clearance (k_e) caratteristiche di ogni tipologia: per esempio per la VWD Vicenza il picco dell'eliminazione è più veloce e spostato verso le 1.5-2.5 h, mentre per i VWD tipo 2B è verso le 4 h e per il soggetto 2A ancora più lento verso le 8 h. Questo aspetto si trascina anche nella successiva analisi di informazione.

Per le diverse stime, infine, si ottengono i valori dei SWR (2.21) i quali vengono riportati per entrambi i modelli nella Tabella 3.9. Poiché i diversi contributi della Φ^{ML} dei termini costanti, dati dal modello di varianza, e variabili, dati dal pool di dati medi disponibili, risultano essere uguali sia per il modello 6M che per il modello 7, si nota, eccetto il caso dei non-0, come le stime del modello 6M presentano valori dei SWR leggermente inferiori, dati dalla diversità dei residui sul fitting. Entrambi comunque sembrano convergere verso lo stesso minimo della funzione obiettivo, tranne per il caso della VWD 2A. Questo perché da Figura 3.9a si osserva la presenza di punti sperimentali di VWF:AG affetti da errori casuali, i quali, non essendo mediati da un pool per la presenza di un solo set di dati, aumentano il valore della funzione obiettivo rispetto al fitting. Quindi il modello 7, pur riducendo il numero di parametri, non migliora di molto la situazione sulla sovrapparametrizzazione.

Tabella 3.9: Valori dei SWR per le stime effettuate con i dati di VWF:AG eCB, riportate nelle tabelle precedenti.

	0	Non-0	VWD Vicenza	VWD 2A	VWD 2B
Modello 6M	45.849	56.549	41.759	118.35	49.338
Modello 7	47.361	56.549	42.245	120.09	50.706

3.5.1 Varianti del modello 6M

Per sondare la struttura del modello 6M, che tra i due modelli proposti meglio descrive nel dettaglio la fase proteolitica, sono state prese in considerazione delle diverse varianti, riassunte in Tabella 3.11. Sono state eseguite delle stime sui dati medi di soggetti normali 0 e non-0, uguagliando dei flussi per comprendere quali parametri risultassero più influenti e critici e inoltre, se siano presenti dei percorsi preferenziali dal flusso F1 al flusso F6 (Tabella 3.2). Le stime

Tabella 3.10: Stime dei parametri del modello 6M e delle varianti analizzate; tra parentesi sono riportati i valori del t-value, quelli indicati in grassetto non superano il test.

Set	Stime parametri															
	variante 1		variante 2		variante 3		variante 4		variante 5		variante 6		variante 7		variante 8	
	0	non-0	0	non-0	0	non-0	0	non-0	0	non-0	0	non-0	0	non-0	0	non-0
k_0	0.0575 (0.01326)	0.1264	0.0604	0.0412 (3.253)	0.0570	0.0413	0.0574	0.0413	0.0606	0.0413	0.0573	0.0412	0.0574	0.0413	0.0607 (6.81e-5)	0.0413
k_1	0.0191 (0.0133)	0.0130	0.0252	0.0077 (4.25e-6)	0.0077	1.22e-6	0.0050	5.06e-7	8.62e-3	9.94e-7	7.89e-3	6.96e-7	1.58e-2	1.17e-7	8.65e-3 (5.53e-8)	1.51e-6
k_2	-	-	-	-	0.0423 (5.73e-6)	0.1124	0.0262	0.0562	0.0474	0.0828	0.0455	0.0832	0.0377	0.0836	0.0476 (3.03e-7)	0.0843
k_3	-	-	0.01003	0.0972 (2.52e-5)	-	-	-	-	0.0047	0.0296	0.0040	0.0297	0.0038	0.0288	0.0044 (9.31e-8)	0.0281
k_4	0.0176 (0.3687)	0.4826	0.8515	2.1176 (4.62e-6)	0.0012	0.0012	0.0019	0.0025	2.87e-4	1.46e-4	8.40e-5	1.46e-4	1.25e-2	1.38e-4	1.20e-3 (1.36e-16)	1.00e-3
k_5	8.38e-5 (0.1757)	1.44e-4	2.78e-4	1.46e-4 (0.7611)	8.35e-5	1.46e-4	8.41e-5	1.46e-4	-	-	-	-	8.42e-5	1.46e-4	2.88e-4 (0.9188)	1.46e-4
k_6	0.0074 (0.2869)	0.0170	1.8230	0.0095 (2.33e-6)	0.0238	0.0436	0.0230	0.0657	-	-	0.0249	0.0429	-	-	-	-
k_e	1.49e-3 (5.252)	6.65e-4	1.43e-3	6.61e-4 (6.851)	1.49e-3	6.61e-4	1.49e-3	6.61e-4	1.43e-3	6.61e-4	1.49e-3	6.61e-4	1.49e-3	6.61e-4	1.43e-3 (5.188)	6.61e-4
D	59.729 (3.07e-6)	52.212	59.847	145.702 (1.36e-6)	61.045	119.494	61.106	111.379	57.12	119.05	60.91	119.24	60.65	118.40	58.24 (9.23e-7)	113.06
t_{max}	33.925 (4.49e-6)	18.612	32.004	32.108 (2.55e-6)	33.872	36.875	33.635	38.589	32.634	36.978	33.721	37.039	33.744	37.108	32.286 (3.68e-6)	38.228
k	1.042 (14.15)	0.955	1.055	0.958 (17.35)	1.042	0.958	1.042	0.958	1.057	0.958	1.042	0.958	1.042	0.958	1.057 (15.72)	0.958
y_b^{CB}	34.982 (3.792)	58.307	32.573	12.197 (4.08e-6)	45.436	89.443	28.517	44.725	54.73	65.87	48.81	65.89	48.98	66.50	54.99 (2.66e-6)	67.05

Tabella 3.11: Varianti analizzate del modello 6M.

Varianti	Parametri uguagliati
variante 1	$k_1 = k_2 = k_3$
variante 2	$k_1 = k_2$
variante 3	$k_1 = k_3$
variante 4	$k_2 = k_3$
variante 5	$k_4 = k_5 = k_6$
variante 6	$k_4 = k_5$
variante 7	$k_4 = k_6$
variante 8	$k_5 = k_6$

parametriche di tutte le varianti prese in considerazione sono riportate in Tabella 3.10.

Il risultato più importante ottenuto dopo l'analisi di tutte le varianti è che i parametri più critici della fase di proteolisi risultano essere k_1 , k_2 e k_3 . Essi rappresentano la produzione dei multimeri UL, H ed L dalle molecole dei SUL. Questa considerazione proviene dai valori delle diverse stime ottenute per ogni variante (Tab. 3.10): per le varianti dalla 1 alla 4 i risultati sono in contrasto con quelli ottenuti col modello 6M semplice, la convergenza dello stimatore converge verso un'altro spazio delle soluzioni portando a dei valori dei SWR molto alti e dei parametri molto distanti. Il comportamento della funzione di verosimiglianza fa convergere il valore dei parametri verso i boundaries imposti e quindi le stime non hanno più significato fisiologico. Questo aspetto invece non accade per le varianti dalla 5 alla 8 che prendono in considerazione la diversa ripartizione dei multimeri. Anzi in queste varianti il comportamento è identico al modello 6M senza variazioni, quindi risultano assumere dei valori conseguenti alla ripartizione dei canali stessi.

3.6 Analisi di informazione e correlazione

Dalle matrici di sensitività $\mathbf{Q}(t)$ (§2 2.3) per ogni risposta si calcola la matrice di informazione $\mathbf{H}(\boldsymbol{\theta})$ (2.11) e si effettua un'analisi d'informazione tramite lo studio dell'andamento della traccia di ogni matrice. I profili delle tracce di $\mathbf{H}(\boldsymbol{\theta})$, sia per la matrice riferita alla risposta di VWF:AG che per quella di VWF:CB, rispecchiano l'andamento del contributo maggiore dato nelle righe di $\mathbf{Q}(t)$, che per tutti i casi di VWF sia per soggetti sani che affetti da disfunzioni è dato dal parametro di eliminazione k_e .

Gli andamenti delle tracce di $\mathbf{H}(\boldsymbol{\theta})$ rispetto alle risposte di VWF:AG, di VWF:CB e della loro somma, vengono riportati in Figura 3.12 per i pazienti VWD Vicenza, per VWD di tipo 2B in

Figura 3.13 mentre per VWD di tipo 2A sono in Figura 3.14. In tutti i casi di VWD nelle figure precedenti, data la criticità della fase proteolitica, vengono riferiti anche i profili delle tracce delle sottomatrici dei parametri di proteolisi (Figure 3.12b, 3.13b, 3.14b). La sottomatrice di proteolisi è quella parte della matrice di informazione $\mathbf{H}(\boldsymbol{\theta})$ riferita ai parametri della proteolisi (k_1, k_2, k_3, k_4, k_5 e k_6).

Questi andamenti ci descrivono l'allocazione ottimale per la conduzione dei nuovi esperimenti, dai quali si potrà ottenere il massimo dell'informazione eseguendo la raccolta dei punti sperimentali negli intervalli temporali a ridosso dei massimi delle curve. Per i profili delle matrici totali, i prelievi sanguigni dovranno essere condotti:

- per VWD Vicenza: 2-4 h;
- per VWD 2B: 6-8 h;
- per VWD 2A: 15-17 h.

Osservando i profili delle tracce riferite alle matrici di proteolisi è prevedibile che i nuovi esperimenti dovranno essere conseguiti nei primissimi istanti iniziali, visto che le dinamiche di tale fase sono molto veloci: i picchi sono quasi tutti prima o nell'intorno delle 2 h per tutti i casi.

Tramite il calcolo della matrice di stimabilità \mathbf{P}_E (2.6) si è calcolata la correlazione tra i parametri. In Tabella 3.12 sono riportate le matrici \mathbf{C} triangolari inferiori sia per il modello 6M che per il 7; esse sono riferite in maniera esemplificativa ai dati medi di VWD 2B ma lo stesso tipo di correlazione, con valori di poco differenti, è presente anche per i restanti altri casi.

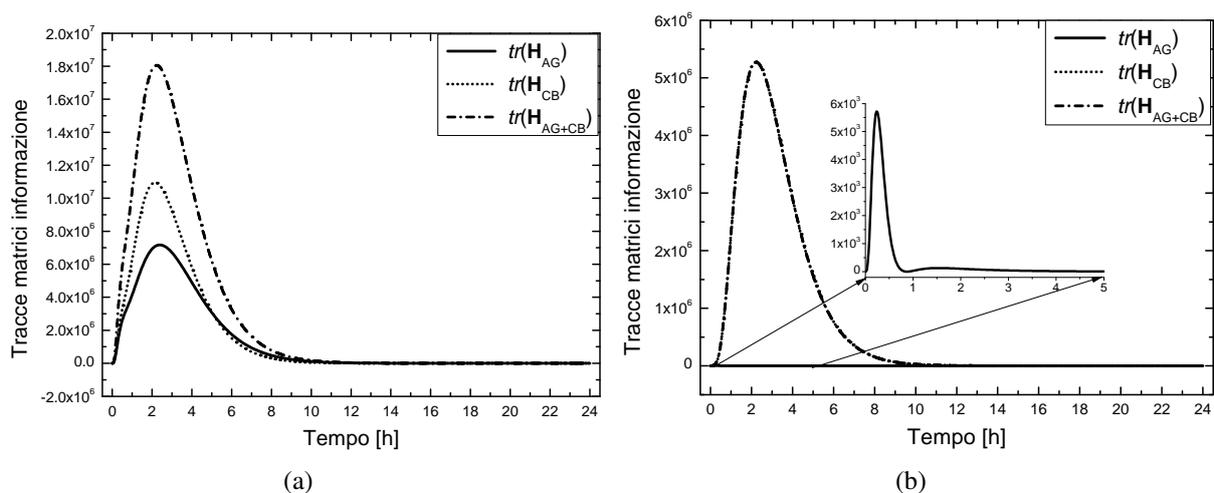


Figura 3.12: Profili delle tracce delle matrici di informazione di antigene e collagene del modello 6M per soggetti affetti da VWD di tipo Vicenza, la figura di sinistra (a) è riferita alla matrice di informazione totale, mentre quella a destra (b) è riferita alla sottomatrice di informazione della fase di proteolisi.

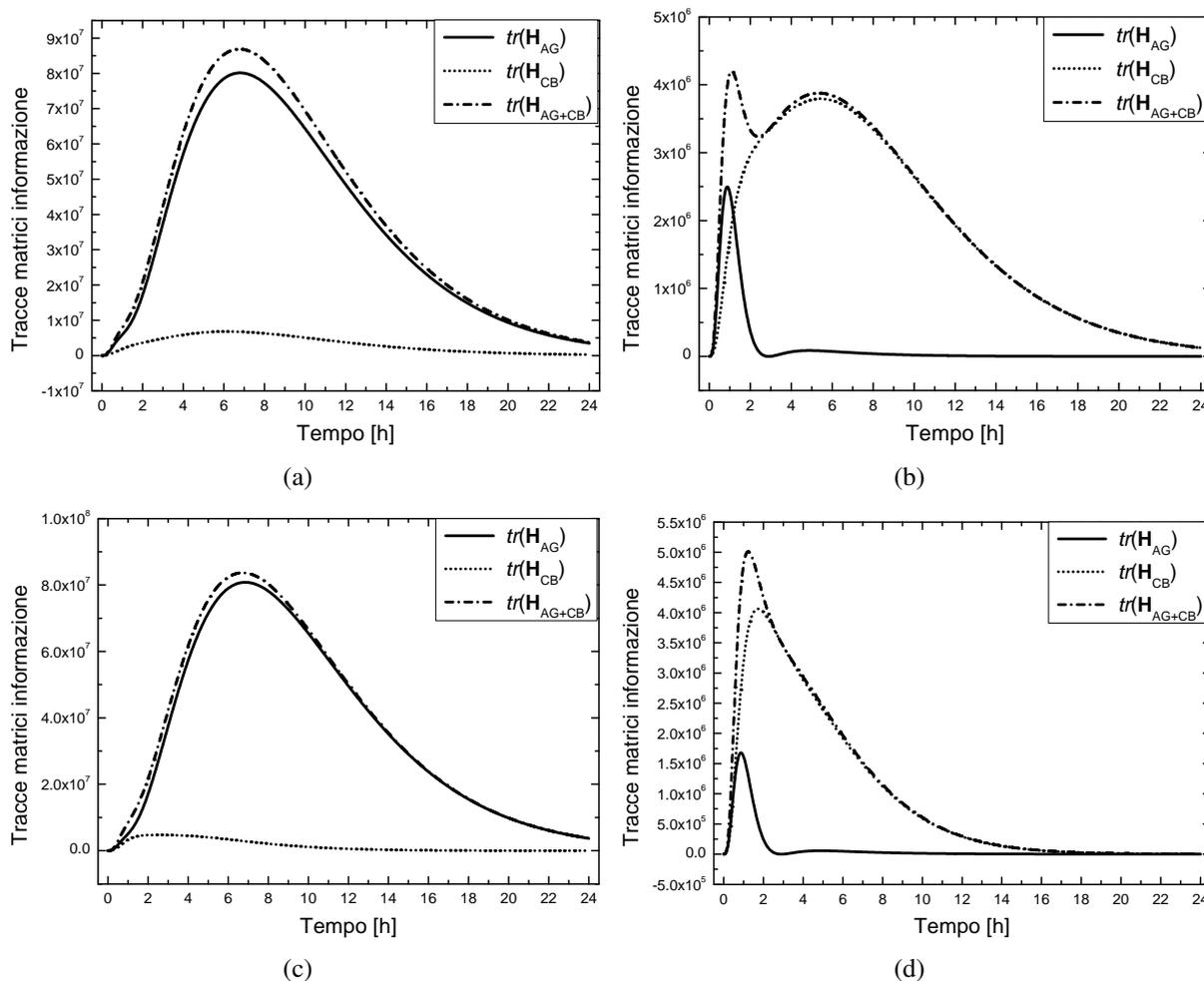


Figura 3.13: Profili delle tracce delle matrici di informazione di antigene e collagene dei modelli 6M (a, b) e 7 (c, d) per soggetti affetti da VWD di tipo 2B, le figure di sinistra sono riferite alla matrice di informazione totale, mentre a destra sono riferite alla sottomatrice di informazione della fase di proteolisi.

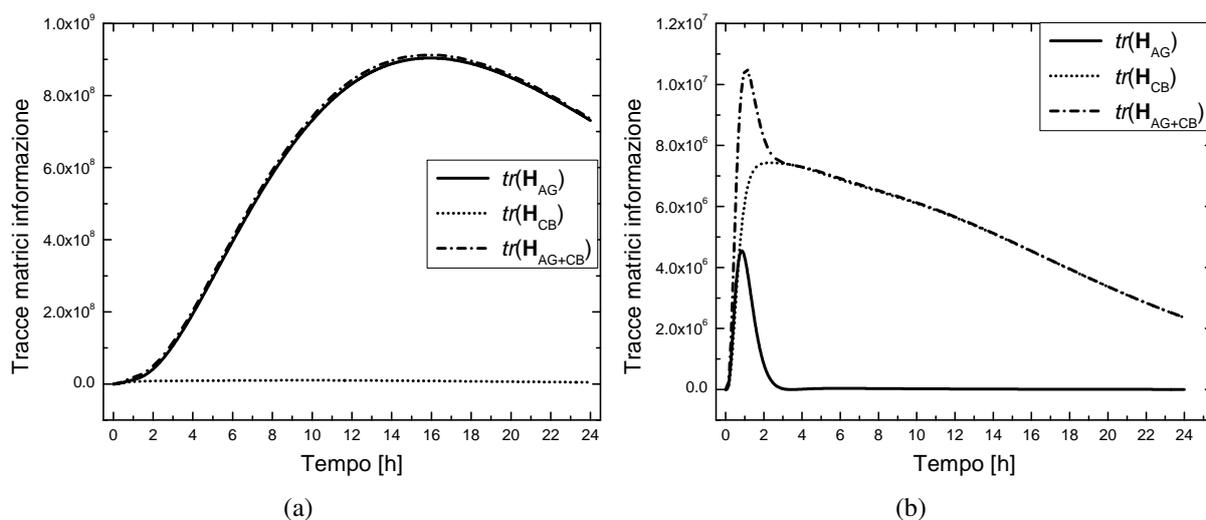


Figura 3.14: Profili delle tracce delle matrici di informazione di antigene e collagene del modello 6M per soggetti affetti da VWD di tipo 2A, la figura di sinistra (a) è riferita alla matrice di informazione totale, mentre quella a destra (b) è riferita alla sottomatrice di informazione della fase di proteolisi.

Tabella 3.12: Matrici di correlazione dei parametri per i modelli 6M e 7, essendo simmetriche vengono riportate solamente le parti triangolari inferiori. Le matrici di correlazione presentano gli stessi parametri correlati, in questo caso sono state calcolate per i parametri rispetto alle stime di dati medi di pazienti affetti da VWD di tipo 2B.

(a) Modello 6M

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e	D	t_{max}	k	y_b^{CB}
k_0	1.000											
k_1	0.352	1.000										
k_2	-0.025	-0.164	1.000									
k_3	0.347	0.984	-0.279	1.000								
k_4	-0.160	-0.534	0.920	-0.619	1.000							
k_5	0.395	0.426	-0.570	0.441	-0.673	1.000						
k_6	0.064	0.229	-0.719	0.395	-0.671	0.221	1.000					
k_e	-0.539	0.103	0.530	0.038	0.415	-0.564	-0.333	1.000				
D	0.999	0.318	-0.040	0.316	-0.160	0.394	0.071	-0.559	1.000			
t_{max}	0.999	0.318	-0.040	0.316	-0.160	0.394	0.071	-0.559	1.000	1.000		
k	-0.346	-0.457	0.843	-0.558	0.896	-0.782	-0.710	0.652	-0.348	-0.348	1.000	
y_b^{CB}	0.144	0.456	-0.915	0.582	-0.951	0.557	0.853	-0.429	0.147	0.147	-0.905	1.000

(b) Modello 7

	k_0	k_1	k_2	k_3	k_e	D	t_{max}	k	y_b^{CB}
k_0	1.000								
k_1	0.009	1.000							
k_2	0.317	-0.233	1.000						
k_3	0.279	-0.625	0.476	1.000					
k_e	-0.572	0.421	0.103	-0.458	1.000				
D	0.999	-0.004	0.288	0.278	-0.589	1.000			
t_{max}	0.999	-0.004	0.288	0.278	-0.589	1.000	1.000		
k	-0.391	0.774	-0.467	-0.792	0.684	-0.393	-0.393	1.000	
y_b^{CB}	0.102	-0.937	0.518	0.652	-0.336	0.105	0.105	-0.822	1.000

Si osserva immediatamente che i tre parametri del rilascio k_0 , D e t_{max} sono altamente correlati tra di loro, mentre D e t_{max} sono completamente correlati. È presente correlazione, seppur in forma minore, tra i parametri della proteolisi k_1 e k_3 e tra k_2 e k_4 . Per sbloccare questa correlazione sarà necessario effettuare una riparametrizzazione dei parametri interessati. L'effetto della correlazione è una conseguenza diretta della sovrapparametrizzazione la quale a sua volta porta a problemi di identificabilità.

Il metodo utilizzato per l'analisi di identificabilità (§2.3.2) è tramite il calcolo del determinante della matrice \mathbf{H} per ogni risposta. Per tutte le tipologie di dati (normali che di VWD), risultano i seguenti valori dei determinanti:

- $\det(\mathbf{H}|_{AG}) \approx 10^{-50} \approx 0$;
- $\det(\mathbf{H}|_{CB}) \approx 10^{30}$;
- $\det(\mathbf{H}|_{AG} + \mathbf{H}|_{CB}) \approx 10^{50}$.

Ciò significa che le misure di antigene non danno alcuna informazione, ed il set di parametri non può essere determinato solamente con le misure di VWF:AG, invece risulta che sia dalle sole misure di collagene che da quelli combinate di antigene e collagene sia possibile identificare univocamente il set di parametri $\hat{\theta}$ del modello.

3.7 Osservazioni sui modelli

Dopo aver analizzato i modelli proposti secondo le tecniche del Capitolo 2, si possono trarre le seguenti osservazioni: i modelli proposti risultano sovrapparametrizzati, le stime convergono verso uno spazio delle soluzioni non univoco; oltretutto non si hanno informazioni sulle statistiche in uscita dallo stimatore poiché la funzione obiettivo presenta un andamento molto piatto vicino al punto di minimo, quindi lo stimatore non è in grado di calcolare la matrice di varianza-covarianza. La fase critica da caratterizzare risulta essere quella di proteolisi poiché determina il riconoscimento delle diverse tipologie e le diverse dinamiche dei flussi. È presente correlazione tra i parametri del rilascio e della proteolisi che assieme alla condizione di sovrapparametrizzazione causano a loro volta un problema di identificabilità. Nel capitolo successivo si cercherà di sbloccare questo aspetto introducendo delle nuove misure di elettroforesi, in primo luogo per osservare il comportamento dei modelli a tali vincoli aggiuntivi e successivamente capire se queste misure danno nuova ed utile informazione.

Capitolo 4

Diagnosi della malattia di von Willebrand mediante modello

Nel presente capitolo si vuole utilizzare il modello 6M come strumento di diagnosi per la VWD. Inizialmente si descrive il metodo utilizzato per effettuare la diagnosi tramite il modello e le modalità di utilizzo dei dati disponibili; successivamente si analizzano le stime parametriche ottenute dalla procedura di diagnosi. Si presenta poi una riparametrizzazione del modello per migliorare l'identificabilità parametrica ed infine si studia l'effetto della variabilità dei valori dei parametri della fase di proteolisi rispetto alle diverse tipologie di campionamento disponibili al fine di diagnosticare correttamente la tipologia della VWD sofferta da un paziente.

4.1 Diagnosi mediante il modello 6M

La VWD presenta un ampio spettro di varianti le quali, come visto nel Capitolo 1, non sempre sono facilmente distinguibili. Effettuare una diagnosi che sia allo stesso tempo attendibile e rapida permette di adottare una terapia efficace e adatta al singolo paziente. Si vuole quindi introdurre il modello matematico della malattia come supporto alla diagnosi.

Avendo a disposizione nuovi dati, oltre a quelli di VWF:AG e VWF:CB, che descrivono le diverse classi multimeriche, si utilizzerà solamente il modello 6M il quale rappresenta il meccanismo della proteolisi con maggior dettaglio nella descrizione dei diversi canali di multimeri. Si abbandona quindi il modello 7 poiché non riesce a caratterizzare la differenza tra i multimeri UL e H. I nuovi dati a disposizione sono delle analisi di elettroforesi individuali in forma di lastre radiografiche, che permettono di ottenere tramite delle analisi d'immagine le diverse frazioni ponderali z_i delle diverse classi di multimeri UL, H ed L. L'estrapolazione delle informazioni sulle frazioni ponderali dalle lastre radiografiche tramite analisi d'immagine viene trattata nel paragrafo successivo.

La procedura che si utilizzerà per diagnosticare la tipologia della VWD del soggetto è illustrata in Figura 4.1. Seguendo la Figura, la procedura inizia partendo da una conoscenza preliminare rappresentata dai dati di frazioni ponderali z_i individuali e dalle stime parametriche sui dati

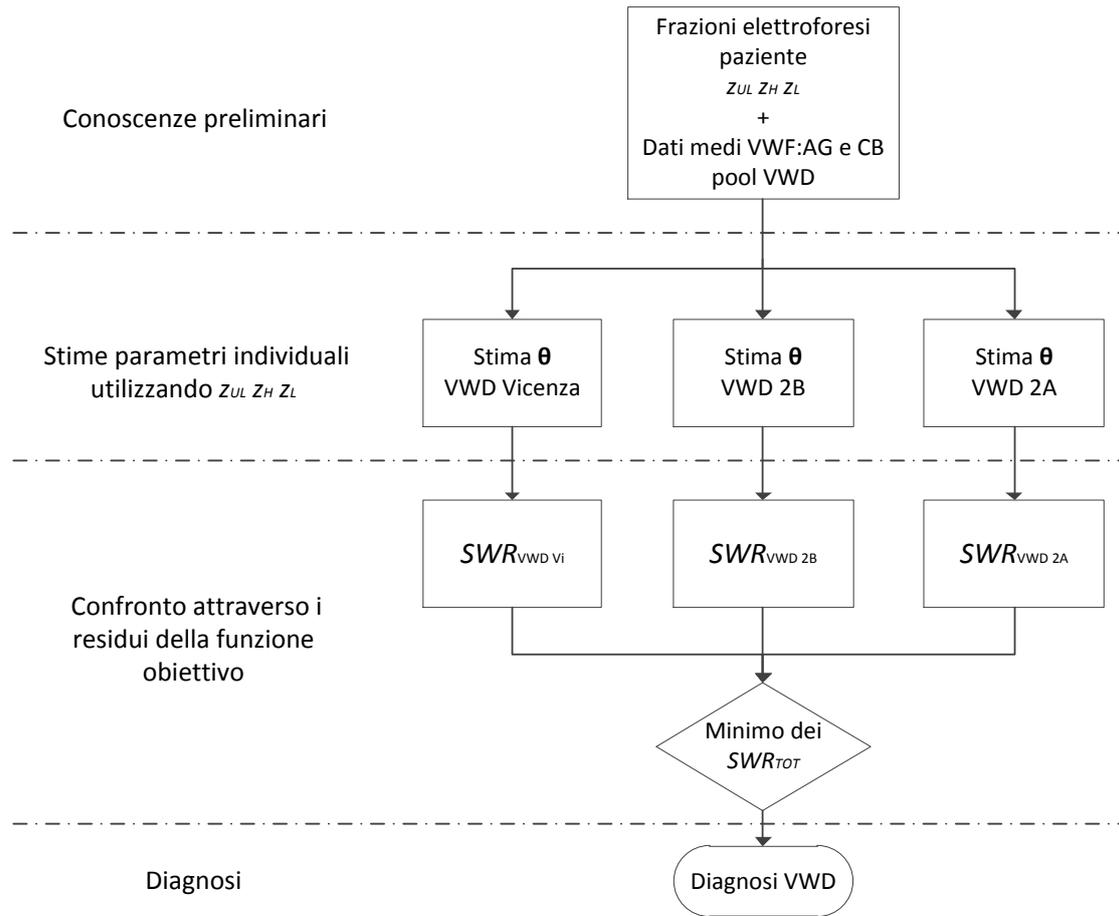


Figura 4.1: Procedura usata in fase di stima per diagnosticare la tipologia di VWD di ogni paziente dai dati di frazioni z_i delle analisi di elettroforesi.

medi di antigene e collagene per ogni pool di VWD; si effettuano successivamente delle stime individuali utilizzando i dati di frazioni ponderali e i dati medi del pool per ogni tipologia (Vicenza, 2B, 2A). Una volta effettuate le tre stime si confrontano i residui della funzione obiettivo Φ^{ML} (eq. 2.16) in uscita dallo stimatore: il fitting che tra le tre stime presenta il valore minimo dei residui totali SWR_{TOT} (eq. 2.21), avrà per quel soggetto e quel pool di VWD la stima dei parametri che più si avvicina al set vero θ individuale. Quindi a tale paziente si potrà diagnosticare la tipologia del pool che ha in uscita dalla stima il valore minimo dei residui.

La funzione obiettivo dello stimatore Φ^{ML} può essere riscritta nel modo seguente:

$$\Phi^{ML} = \beta + \sum_i SWR_i \quad i = y_{AG}, y_{CB}, z_{UL}, z_H, z_L \quad (4.1)$$

la quale è data dalla somma di un termine costante β che dipende dal modello di varianza scelto per l' i -esima risposta, dai punti sperimentali e dalle componenti variabili della matrice di

varianza-covarianza sulle misure della i -esima risposta, e dalla sommatoria dei residui pesati SWR_i calcolati dalle differenze tra le misure sperimentali e i punti predetti dalla stima parametrica della i -esima risposta. Si possono quindi riportare, nei risultati successivi, i contributi di Φ^{ML} in tre distinte parti:

- costanti (modello di varianza);
- variabili (misure sperimentali);
- SWR (residui sulle risposte predette).

È importante inoltre separare l'analisi dei SWR per ogni i -esima risposta in aggiunta alla somma totale. Così facendo si separa l'informazione ottenibile dal fitting di ogni risposta per i reciproci dati sperimentali, osservando in che modo la stima parametrica riesca a rappresentare le frazioni ponderali individuali z_i , o se i profili predetti di VWF:AG e VWF:CB rimangono all'interno della banda di variabilità del pool; questo aspetto lo si vedrà nel dettaglio nel §4.3 e §4.4.

Tra i dati di elettroforesi disponibili sono stati presi in considerazione quattro pazienti esemplificativi:

- i pazienti A e B, dei quali non si conosce la diagnosi tramite antigene e collagene;
- i pazienti G e M i quali vengono diagnosticati come VWD tipo Vicenza tramite i dati individuali di antigene e collagene, i quali fanno parte del pool VWD Vicenza ¹.

Successivamente la procedura di Figura 4.1 verrà applicata anche utilizzando i dati di VWF:AG e CB individuali per i pazienti G ed M assieme alle frazioni ponderali e i dati medi dei pool.

4.2 Dati da analisi di gel-elettroforesi

Come descritto nel §1.5, le radiografie rappresentano qualitativamente la distribuzione dei diversi multimeri lisi dall'enzima ADAMTS13 nei diversi istanti temporali. Avendo a disposizione un nuovo tipo di informazione si è pensato di tradurre queste analisi qualitative in dati quantitativi in termini di frazioni di multimeri, o concentrazioni, da aggiungere in fase di stima ai dati di antigene e collagene medi. Capire se i dati di elettroforesi possano descrivere nel modo migliore la fase di proteolisi è uno dei punti cardine: cioè se tale informazione aggiuntiva riesce a caratterizzare i canali multimerici rispetto ai soli dati di VWF:AG e VWF:CB, costringendo il modello verso la convergenza di uno spazio delle soluzioni dei parametri critici vicino al set vero θ , riducendo l'incertezza sulla stima del set stesso.

¹Purtroppo non è noto se i dati di VWF:AG e CB corrispondano allo stesso campionamento per le analisi di gel-elettroforesi o siano stati effettuati in periodi diversi a distanza di anni

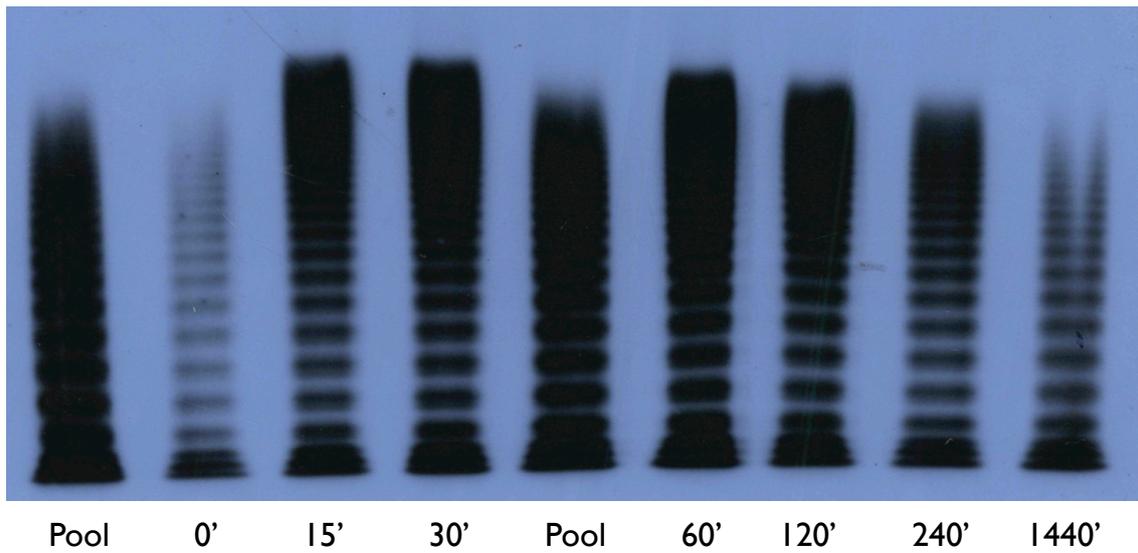


Figura 4.2: Esempio di lastra radiografica delle analisi di gel-elettroforesi per il paziente B.

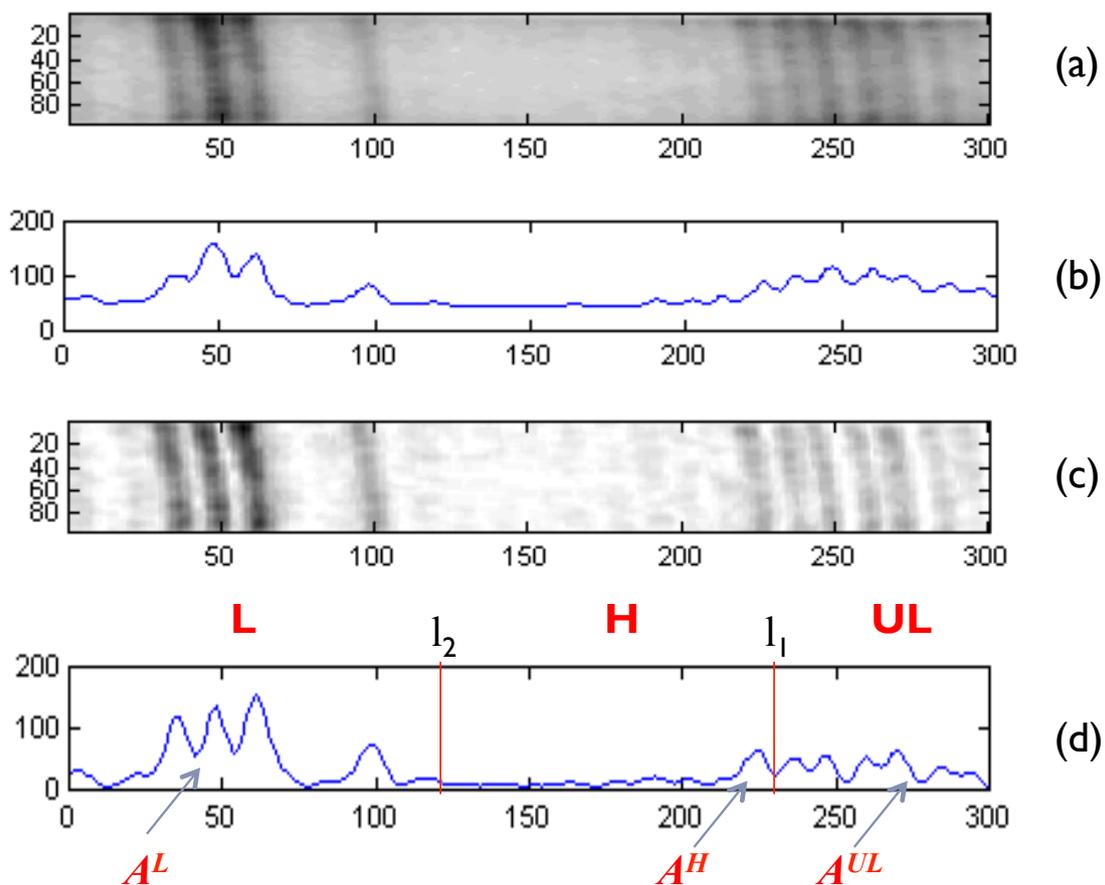


Figura 4.3: Fasi dell'analisi di immagine per estrapolare i profili di densità ottica AUC. Le varie fasi sono nell'ordine: (a) pattern multimerico da analizzare ridotto in scala di grigi, (b) profilo di densità ottica corrispondente, (c) rimozione dello sfondo, (d) nuovi profili di densità ottica e corrispettivi intervalli dei diversi multimeri e aree associate del pattern (Ye et al., 1999).

Un esempio di lastra radiografica di un paziente, tra i nove disponibili, è riportato in Figura 4.2. Ogni striscia o pattern multimerico rappresenta un campionamento sperimentale ad un determinato periodo dall'inizio dell'esperimento, mentre i pattern chiamati pool si riferiscono a degli standard di soggetti normali in condizioni basali e fungono da riferimento.

Tramite delle metodologie di analisi d'immagine applicate all'elettroforesi (Ye *et al.*, 1999), i pattern multimerici vengono ridotti in curve di aree di densità ottica per ogni istante temporale (Figura 4.3a). Si stabilisce a priori dai pool di riferimento una soglia di distinzione tra le bande dei multimeri UL ed H e tra H ed L, sapendo che in condizioni basali a $t = 0$ i multimeri UL non sono presenti in un soggetto normale (Casonato *et al.*, 2010). Le frazioni ponderali delle singole unità multimeriche z_{UL} , z_H e z_L di un pattern vengono calcolate tramite i rapporti delle aree di intensità ottica secondo l'equazione:

$$z_i|_{t_{sp}} = \frac{A_i}{\sum_{j=1}^{N_z} A_j} \quad (4.2)$$

dove A_i sono le aree sotto la curva di intensità ottica corrispettive alle singole frazioni z_i di un pattern e N_z il numero di frazioni, pari a 3 (z_{UL} , z_H , z_L). Effettuando tale operazione per ciascun pattern ad ogni tempo di campionamento si costruiscono i profili temporali delle frazioni (Galvanin, 2012d), un esempio è rappresentato in Figura 4.4.

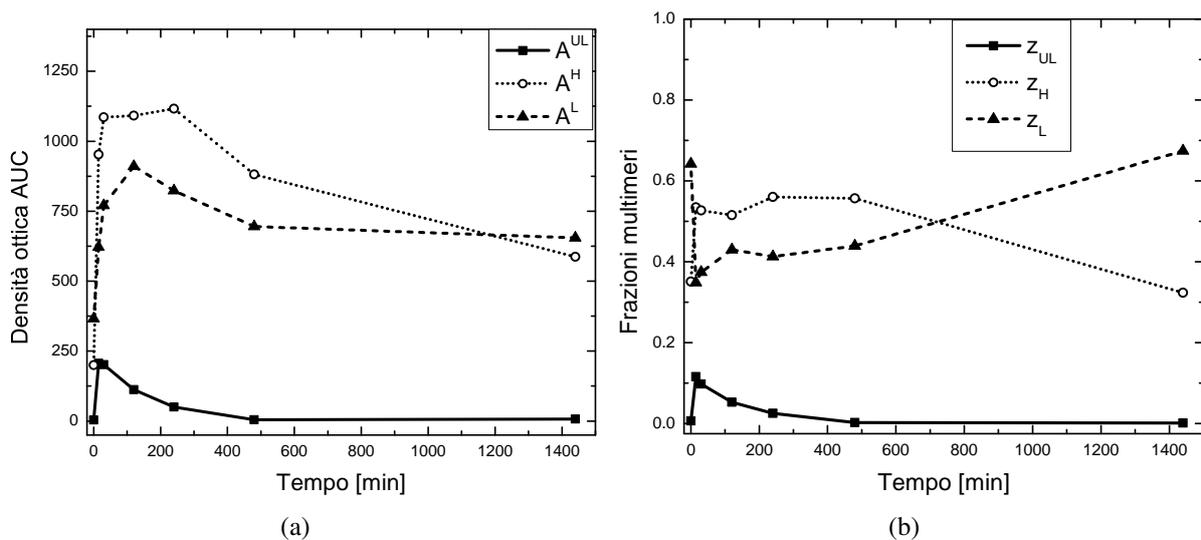


Figura 4.4: Esempio di punti di densità ottica estrapolati dalle lastre di elettroforesi (a) e rispettivi punti in termini di frazioni calcolati secondo (4.2), le linee hanno solamente lo scopo di congiungere i punti sperimentali ed aiutare la rappresentazione grafica.

4.3 Utilizzo dei dati disponibili

A questo punto si hanno a disposizione delle informazioni aggiuntive in termini di frazioni multimeriche per alcuni pazienti e sorge il problema di come inserire tali dati in fase di stima. Le stime parametriche del Capitolo 3 sono state condotte considerando i dati disponibili di VWF:AG e VWF:CB rappresentativi di un paziente medio di ogni classe di dati, dunque il modello di varianza utilizzato è riferito al sistema di misurazione, con deviazione standard costante e pari a $\omega = 3$ U/dL (§2.4). Nel caso in cui si voglia effettuare una stima associando i dati di elettroforesi di un singolo paziente ai dati di antigene e collagene medi del pool di pazienti, si dovrà utilizzare:

- per i dati di frazioni un modello a varianza costante, caratteristico dell'errore commesso nel valutare le frazioni;
- per i dati del pool un modello di varianza eteroschedastico (§2.4) che consideri la dipendenza della misura dalla variabilità intra-individuale.

La stima condotta secondo tali modalità ha un significato molto importante poiché si vuole stimare il set $\hat{\theta}$ caratteristico del singolo, confrontandolo con lo spazio delle soluzioni del pool di ogni tipologia VWD rappresentato appunto da un valore medio ed una banda di variabilità.

Per le stime seguenti di questo capitolo si è utilizzato per le frazioni z_i un valore di deviazione standard costante pari a $\omega = 0.05$, mentre per il modello eteroschedastico sono stati stimati con il metodo dei minimi quadrati e per ogni tipologia di dati, i parametri di varianza ω e γ dai valori forniti di deviazione standard delle misure del pool rispetto al valore medio calcolato. I parametri di varianza eteroschedastica stimati sono riportati nella Tabella 4.1. Per quanto riguarda la tipologia VWD 2A, non avendo un pool di pazienti, si sono utilizzati i parametri dei 2B assumendo la stessa variabilità intra-individuale.

Tabella 4.1: Stime dei parametri di varianza eteroschedastica ω [U/dL] e γ per ogni pool di dati.

		0	Non-0	VWD Vicenza	VWD 2B
VWF:AG	ω	0.63	5.51	0.59	0.48
	γ	0.50	0.18	0.66	0.67
VWF:CB	ω	0.28	2.55	0.24	0.20
	γ	0.69	0.34	0.89	1.08

4.4 Risultati delle stime per la diagnosi della VWD

Nella Tabella 4.2 vengono riportati i valori dei residui SWR_i di ogni stima effettuata per ciascuno dei quattro pazienti presi in esame. Tutte le stime parametriche effettuate di ogni paziente e per ciascun caso sono riportate in Appendice A. Sono stati utilizzati per gli intervalli di stima dei limiti molto ampi (0-100 in termini normalizzati) in modo tale da assicurare una variabilità ampia dei parametri, tranne per i parametri relativi alla correzione k ed y_b^{CB} dove i limiti corrispondono a quelli utilizzati nelle stime del Capitolo 3.

Confrontando i valori dei diversi SWR si osserva che non sempre i valori minimi dei residui, evidenziati in grassetto nella Tabella 4.2, di ogni paziente sono riferiti alla stessa tipologia.

Osservando i residui per i pazienti G ed M i valori minimi dei SWR_{pool} sono in corrispondenza della tipologia VWD Vicenza: i valori minimi dei SWR_{fraz} per G sono riferiti alla VWD tipo Vicenza; per M invece sono in corrispondenza della 2B, anche se la differenza con i SWR_{fraz} per la VWD Vicenza è minima (0.58). Dunque si può diagnosticare per i pazienti G ed M una VWD di tipo Vicenza, confermando la diagnosi del personale medico. Nello specifico per il paziente G si possono notare graficamente in Figura 4.5 gli esiti delle stime per la diagnosi sul fitting di VWF:AG e VWF:CB secondo la procedura di Figura 4.1, i valori delle tre stime sono riportate in sintesi nella Tabella 4.3. È facile constatare che il fitting migliore risulta quello con il pool VWD Vicenza (Figura 4.5a,b), mentre per il pool VWD 2B si ha solamente un buon fitting per il VWF:AG dove la curva rimane all'interno della banda di variabilità (Figura 4.5c), ma non per il VWF:CB dove la curva tende ad assumere dei valori inferiori (Figura 4.5d). La stima associata al pool VWD 2A manifesta l'esito negativo del test per l'antigene, per il collagene invece la curva rimane all'interno della banda di variabilità (Figura 4.5e,f), ma se si osservano i diversi profili predetti sul fitting delle frazioni multimeriche rispetto ai punti sperimentali ottenuti dalle lastre di elettroforesi, riportati in Figura 4.7a,c,e sempre per le tre stime effettuate, si osserva come il fitting sia migliore per la VWD Vicenza e, pur con dei residui maggiori, anche per la 2B (Figura 4.7a,c), invece per la tipologia VWD 2A non ci sia una buona predizione delle condizioni basali e finali (Figura 4.7e).

Per ogni stima si possono notare gli effetti sui parametri in Tabella 4.3, dove per alcuni canali proteolitici i parametri hanno un valore nullo e quindi il modello riconosce delle anomalie in alcune tipologie. Per il parametro di eliminazione (k_e) non sono presenti degli scostamenti dalle stime preliminari, ma sono presenti dei valori anomali dei parametri del rilascio (k_0 , D , t_{max}) i

Tabella 4.2: Valori dei SWR per ciascuna stima secondo la procedura in Figura 4.1: i valori minimi sono evidenziati in grassetto. I diversi contributi sono riportati separatamente in: termini costanti relativi al modello di varianza scelto, termini variabili in riferimento ai dati sperimentali a disposizione e termini residui sulla differenza tra i profili predetti dalla stima e i punti sperimentali. I termini *fraz* e *pool* rappresentano la somma rispettivamente dei termini delle frazioni e di YWF:AG/CB del pool.

Paziente e dati	VWD Vicenza		SWR	VWD 2B		SWR	VWD 2A		SWR	
	Costanti	Variabili		Costanti	Variabili		Costanti	Variabili		
Rif tot	49.15			56.55			59.89			
A	ZUL	5.514	-17.97	0.53	5.514	-17.97	0.158	5.514	-17.97	1.829
	ZH	5.514	-17.97	0.388	5.514	-17.97	0.054	5.514	-17.97	36.42
	ZL	5.514	-17.97	0.552	5.514	-17.97	0.219	5.514	-17.97	42.67
	<i>fraz</i>			-35.91			-36.95			43.54
	YAG:pool YCB:pool pool	9.189 9.189	16.75 13.3	0.789 1.684	9.189 9.189	21.73 15.21	1.071 0.657	9.189 9.189	26.66 11.65	8.526 14.05
B	ZUL	6.433	-20.97	1.234	6.433	-20.97	0.286	6.433	-20.97	4.606
	ZH	6.433	-20.97	5.063	6.433	-20.97	3.152	6.433	-20.97	23.52
	ZL	6.433	-20.97	4.377	6.433	-20.97	3.405	6.433	-20.97	35.8
	<i>fraz</i>			-32.94			-36.77			20.32
	YAG:pool YCB:pool pool	9.189 9.189	16.75 13.3	0.127 1.848	9.189 9.189	21.73 15.21	2.378 2.378	9.189 9.189	26.66 11.65	23.5 30.58
G	ZUL	6.433	-20.97	1.566	6.433	-20.97	0.901	6.433	-20.97	1.113
	ZH	6.433	-20.97	0.693	6.433	-20.97	2.672	6.433	-20.97	16.44
	ZL	6.433	-20.97	0.674	6.433	-20.97	2.313	6.433	-20.97	22.48
	<i>fraz</i>			-40.68			-37.73		0	-3.585
	YAG:pool YCB:pool pool	9.189 9.189	16.75 13.3	0.216 0.969	9.189 9.189	21.73 15.21	3.345 2.224	9.189 9.189	26.66 11.65	15.31 13.09
M	ZUL	6.433	-20.97	2.598	6.433	-20.97	1.824	6.433	-20.97	1.643
	ZH	6.433	-20.97	1.364	6.433	-20.97	1.796	6.433	-20.97	7.705
	ZL	6.433	-20.97	0.821	6.433	-20.97	0.579	6.433	-20.97	14.57
	<i>fraz</i>			-38.83			-39.41			-19.69
	YAG:pool YCB:pool pool	9.189 9.189	16.75 13.3	0.342 1.176	9.189 9.189	21.73 15.21	1.449 2.112	9.189 9.189	26.66 11.65	5.767 9.548
			49.95			58.89				

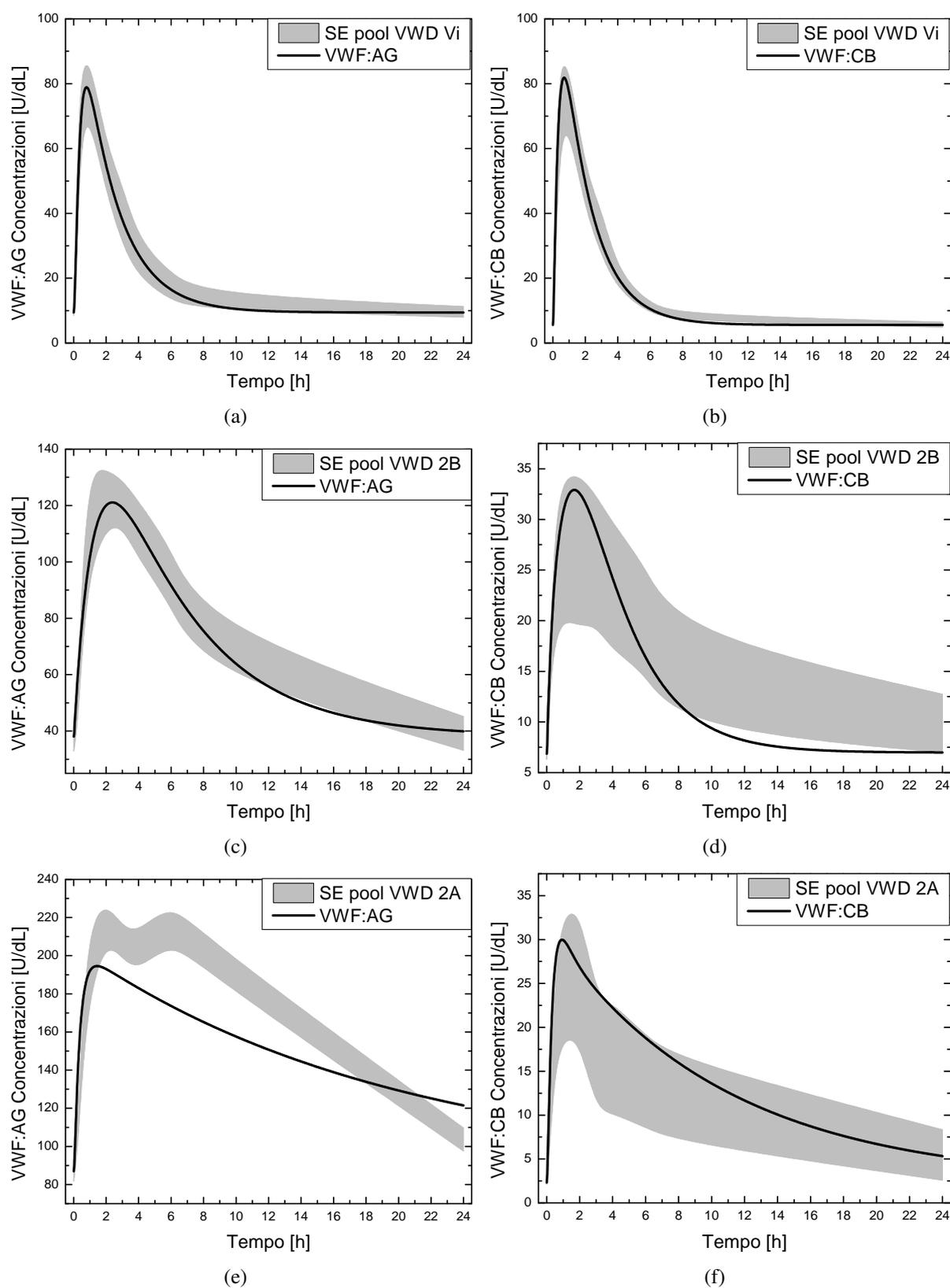


Figura 4.5: Stime effettuate per il paziente G in base alla procedura per la diagnosi mediante modello. Le curve continue rappresentano i profili di VWF:AG e VWF:CB predetti dalla stima associata al pool Vicenza (a,b), per il pool 2B (c,d) e 2A (e,f); le aree grigie invece rappresentano la variabilità di ogni pool utilizzando i modelli di varianza del §4.2.

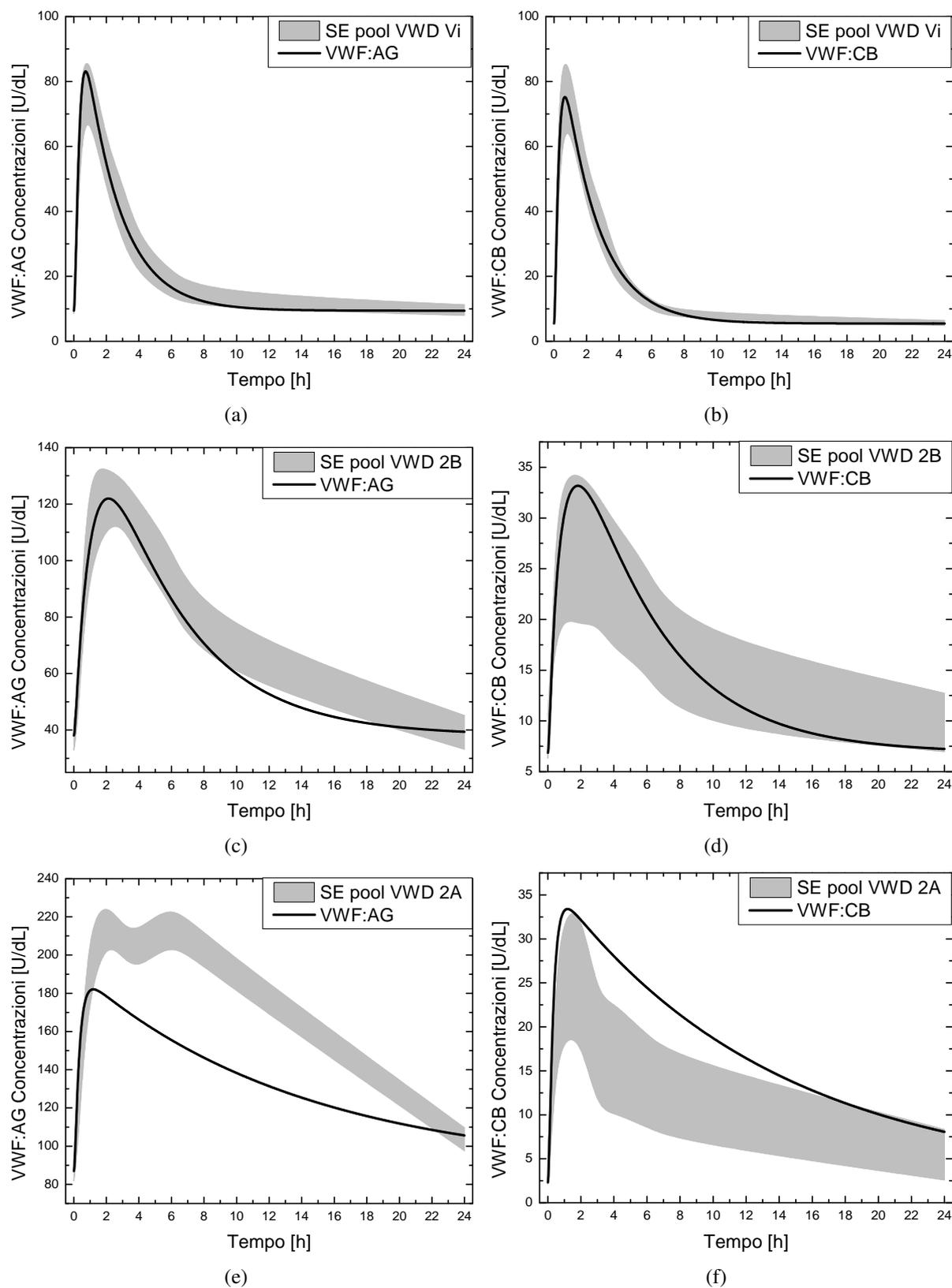


Figura 4.6: Stime effettuate per il paziente B in base alla procedura per la diagnosi mediante modello. Le curve continue rappresentano i profili di VWF:AG e VWF:CB predetti dalla stima associata al pool Vicenza (a,b), per il pool 2B (c,d) e 2A (e,f); le aree grigie invece rappresentano la variabilità di ogni pool utilizzando i modelli di varianza del §4.2.

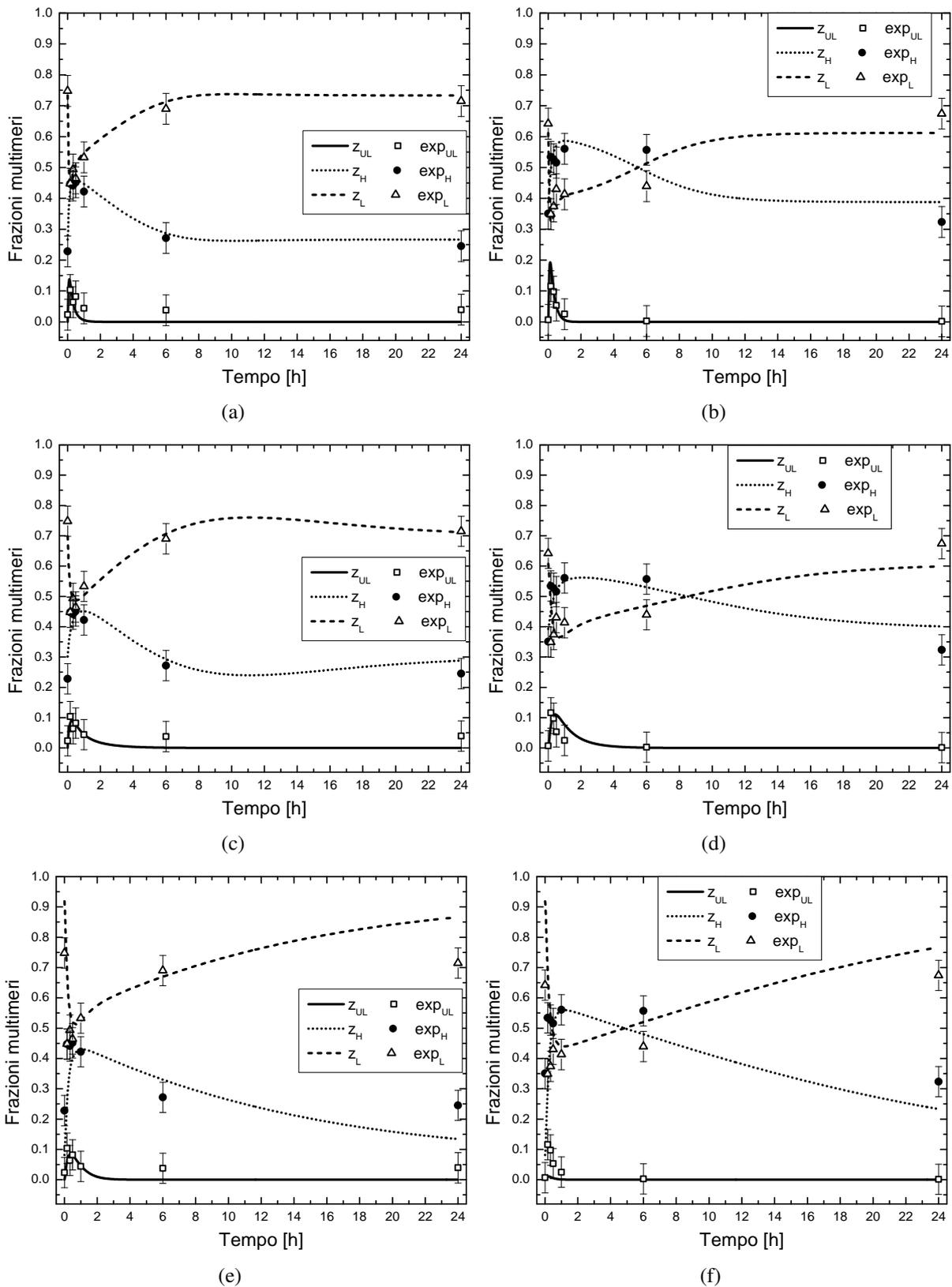


Figura 4.7: Profili di frazioni predetti da ogni stima associata per il paziente G, le figure a sinistra, ed il paziente B, le figure alla destra in riferimento alle stime di Figura 4.1; (a, b) VWD Vicenza, (c, d) VWD 2B, (e, f) VWD 2A.

quali assumono valori distanti dalle stime di riferimento θ_0 , molto probabilmente per aggiustare il fitting sui dati sperimentali qualora il pool di riferimento non sia quello diagnosticato dal paziente G.

Tabella 4.3: Stime parametri per la diagnosi della VWD del paziente G; i valori θ_0^i sono i fattori di normalizzazione dati dalle stime del Capitolo 3.

	VWD Vicenza	θ_0^{Vi}	VWD 2B	θ_0^{2B}	VWD 2A	θ_0^{2A}
k_0	0.0593	0.0584	3.319	0.078	0.3035	0.0975
k_1	0.07842	0.0351	0.005456	0.0167	0.01306	0.0064
k_2	0.0775	0.154	0.007876	0.00509	0.03703	0.0014
k_3	1.492×10^{-8}	3.18×10^{-9}	0	0.0023	0	0.0172
k_4	0	0.0236	0	0.00439	0	0.0026
k_5	0.001372	0.00136	0.002763	5.54×10^{-5}	0.0007201	0.00037
k_6	0.3508	0.0887	0.07843	0.15	0.0304	0.0155
k_e	0.007778	0.00789	0.00314	0.0028	0.000853	0.0011
D	183.2	59.7	1.052	44.98	9.197	35.92
t_{\max}	12.41	32.25	1.868	30.39	12.9	29.32
k	0.5911	0.719	0.1832	0.229	0.02664	0.0444
y_b^{CB}	2.515	5.58	11.46	10.05	7.074	3.93

Tabella 4.4: Stime parametri per la diagnosi della VWD del paziente B; i valori θ_0^i sono i fattori di normalizzazione dati dalle stime del Capitolo 3.

	VWD Vicenza	θ_0^{Vi}	VWD 2B	θ_0^{2B}	VWD 2A	θ_0^{2A}
k_0	0.09644	0.0584	0.2542	0.078	0.2192	0.0975
k_1	0.0559	0.0351	0.005788	0.0167	0.004428	0.0064
k_2	0.0432	0.154	0.01024	0.00509	0.05532	0.0014
k_3	9.017×10^{-8}	3.18×10^{-9}	0	0.0023	0	0.0172
k_4	0.0804	0.0236	0	0.00439	0.1607	0.0026
k_5	0	0.00136	0	5.54×10^{-5}	4.241×10^{-5}	0.00037
k_6	0.1719	0.0887	0.04242	0.15	3.279×10^{-6}	0.0155
k_e	0.007688	0.00789	0.003336	0.0028	0.001207	0.0011
D	29.45	59.7	15.44	44.98	23.52	35.92
t_{\max}	26.72	32.25	13.74	30.39	13.07	29.32
k	0.582	0.719	0.1832	0.229	0.02664	0.0444
y_b^{CB}	3.656	5.58	14.94	10.05	7.074	3.93

Le stesse considerazioni del paziente G si possono compiere sul paziente M, per sintesi si omette di riportare i grafici delle stime. Essendo disponibili i dati di VWF:AG e VWF:CB individuali di G ed M successivamente si aggiungono nelle stime per osservare il cambiamento con tale informazione aggiuntiva.

I casi critici si hanno per i pazienti A e B: i valori minimi dei SWR_{fraz} sono in corrispondenza della tipologia 2B, mentre i valori minimi dei SWR_{pool} per la VWD tipo Vicenza. Dunque se si dovesse diagnosticare la VWD sulla base del fitting sulle frazioni di multimeri i pazienti A e

B sembrano soffrire della VWD tipo 2B, mentre sul fitting di VWF:AG e VWF:CB per i pool la diagnosi propende per una VWD Vicenza. In Figura 4.6 vengono riportati gli esiti delle tre stime del paziente B. effettuate per ciascun pool, rispetto alle stime parametriche effettuate e riportate sinteticamente in Tabella 4.4. Il fitting migliore risulta essere quello con il pool VWD Vicenza (Figura 4.6a,b), mentre per il pool VWD 2B si ha solamente un buon fitting per il VWF:CB dove la curva rimane all'interno della banda di variabilità (Figura 4.6d) Per il pool VWD 2A invece non si ha un buon fitting sia per l'antigene che per il collagene (Figura 4.6e,f). Analizzando invece il fitting sulle frazioni ponderali multimeriche, riportati in Figura 4.7b,d,f, si nota come il fitting migliore risulta quello per la VWD 2B dove i profili rientrano nelle SE dei punti sperimentali soprattutto per i multimeri UL, ma si ha una cattiva predizione delle condizioni finali per gli H ed L; il fitting sembra buono anche per la stima con il pool Vicenza visto che i SWR_{fraz} sono in linea con quelli sulle stime per i 2B. Il fitting sul pool VWD 2A invece non è buono nemmeno per le frazioni multimeriche e si può escludere con certezza questa tipologia (Figura 4.7f).

Per quanto riguarda gli effetti sui parametri del paziente B di ogni stima per la diagnosi (Tabella 4.4) si possono notare, anche in questo caso, che alcuni canali presentano parametri della fase proteolitica con valore nullo, mentre si osservano gli stessi effetti sui parametri del rilascio visti per il paziente G.

La diagnosi per B è più affine alla VWD Vicenza osservando le stime sul pool di VWF:AG e CB, ma sulle frazioni sembra essere più incerta e propendere per la VWD 2B. Una spiegazione a tal proposito la si può trovare secondo due ipotesi: la prima di carattere numerico è data dalla maggior banda di variabilità (in termini di deviazione standard) della tipologia 2B che influenza la convergenza della funzione obiettivo, ricordando la dipendenza dalla matrice di varianza (eq. 2.16); la seconda ipotesi è più a monte, nella mancanza del pool per la VWD di tipo 1 e quindi di un confronto dei SWR_i con il pool di tale tipologia. Gli andamenti della VWD 1 presentano caratteristiche simili di VWF:AG e CB ad un paziente normale ma con valori quantitativamente inferiori, caratteristica più vicina alla tipologia VWD Vicenza ma con una clearance più lenta. Per il paziente A (si omette di riportare i grafici per sintesi) si può diagnosticare con più certezza una VWD Vicenza, essendoci una più marcata differenza dei valori SWR_{pool} tra VWD Vicenza e 2B, ma comunque permane una differenza minore tra i valori dei SWR_{fraz} (Tabella 4.2).

Concludendo si può affermare che le stime dei dati di elettroforesi associati ai pool riescono a distinguere con certezza la tipologia 2A dalle tipologie Vicenza e 2B. Non è possibile trarre

una valutazione generale data la mancanza di ulteriori dati sui pazienti 2A, vista la rarità della tipologia si è assunta una variabilità del pool simile alla tipologia 2B e non è detto che questa ipotesi sia corretta. Inoltre la totale mancanza di dati sulla tipologia VWD 1 non permette il confronto sul fitting con le altre tipologie, soprattutto con la tipologia Vicenza poiché essendo molto affine (§1.5) il modello potrebbe riconoscere dei pazienti VWD Vicenza quando essi in realtà sono affetti da VWD tipo 1.

4.4.1 Diagnosi da dati VWF:AG e CB individuali

Avendo a disposizione i dati singoli di VWF:AG e VWF:CB per i pazienti G e M si è pensato di inserire tale informazione nelle stime parametriche associandoli alle frazioni di elettroforesi individuali e ad ogni pool. Le tre stime parametriche di ogni paziente sono riportate nel dettaglio in Appendice A, mentre i valori dei SWR_i di tali stime sono presenti in Tabella 4.6.

Tabella 4.5: Stime parametri per la diagnosi con dati dei singoli pazienti di VWF:AG e CB, per il paziente G (a) ed il paziente M (b).

(a) Stime paziente G						
	Stime	Valori normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.0712	1.219	0.0584	1	0	100
k_1	0.2067	5.888	0.0351	1	0	100
k_2	0.03425	0.222	0.154	1	0	100
k_3	4.42×10^{-8}	13.9	3.18×10^{-9}	1	0	100
k_4	0.3518	14.91	0.0236	1	0	100
k_5	0.00116	0.851	0.00136	1	0	100
k_6	0.4657	5.25	0.0887	1	0	100
k_e	0.00979	1.241	0.00789	1	0	100
D	59.89	1.003	59.7	1	0	100
t_{\max}	31.12	0.965	32.25	1	0	100
k	0.64	0.885	0.719	1	0.6	1.5
y_b^{CB}	2.47	0.442	5.58	1	0.2	1.8

(b) Stime paziente M						
	Stime	Valori normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.0093	0.159	0.0584	1	0	100
k_1	0.03025	0.862	0.0351	1	0	100
k_2	0.03348	0.217	0.154	1	0	100
k_3	8.5×10^{-9}	2.673	3.18×10^{-9}	1	0	100
k_4	0	0	0.0236	1	0	100
k_5	0.00132	0.973	0.00136	1	0	100
k_6	0.1715	1.934	0.0887	1	0	100
k_e	0.01884	2.388	0.00789	1	0	100
D	286.9	4.806	59.7	1	0	100
t_{\max}	65.67	2.036	32.25	1	0	100
k	0.57	0.796	0.719	1	0.6	1.5
y_b^{CB}	2.63	0.472	5.58	1	0.2	1.8

Osservando la Tabella 4.6 è evidente che questa ulteriore informazione porta ad una diagnosi di VWD Vicenza sicura per entrambi i pazienti G ed M: si notino i valori elevati sia dei residui totali (SWR_{pool}) che parziali per antigeni e collagene ($SWR_{AG,pool}$, $SWR_{CB,pool}$) delle stime per i pool di VWD 2B e 2A, rispetto ai residui per le stime del pool VWD Vicenza. Si osserva in questo caso, rispetto alla Tabella 4.2, che i valori per i residui sulle frazioni multimeriche (SWR_{fraz}) sono nettamente minori per la VWD Vicenza. Ciò significa che utilizzando congiuntamente i dati individuali di VWF:AG e CB alle frazioni ponderali e ai dati medi del pool la diagnosi può essere fatta con più sicurezza.

Questo aspetto lo si nota nelle stime parametriche per ogni pool: le stime con i dati medi di VWD 2B ed VWD 2A assumono valori privi di significato, a causa dell'elevato valore di Φ^{ML} , quindi non si riesce ad arrivare a convergenza di una soluzione del set parametrico per gli intervalli di stima imposti. Si riportano in Tabella 4.5 le stime ottenute aggiungendo i dati individuali solo per i casi associati al pool VWD Vicenza.

4.5 Stima dei parametri individuali

Un aspetto interessante sull'identificabilità del set parametrico lo si nota esaminando le stime condotte utilizzando solamente i dati individuali, sia di VWF:AG/CB che per le frazioni z_i . Sono state effettuate le stime per i pazienti G ed M: un aspetto importante è dato dal fatto che per il paziente G si ottengono dallo stimatore le statistiche, ciò significa che la funzione obiettivo presenta un minimo globale in quell'intorno dello spazio delle soluzioni e il set parametrico da sicuramente non identificabile diventa probabilmente identificabile.

Nella Tabella 4.7 è riportata la stima condotta solamente con i dati individuali ed i valori dei residui per il paziente M, mentre nella Tabella 4.8 sono riportati i valori della stima per il paziente G. Per un confronto vengono affiancate le stime del paragrafo precedente associate al pool. I parametri stimati hanno dei valori in linea con le stime effettuate in precedenza associate al pool (Tabella 4.7 colonna *Stime pool*), dove k_4 assume un valore nullo, mentre si notino i valori molto elevati del parametro D (≈ 300).

Anche per il paziente G i parametri stimati hanno dei valori in linea con le stime effettuate in precedenza associate al pool (Tabella 4.8 colonna *Stime pool*) tranne per k_2 il quale assume un valore doppio. Essendo presenti le statistiche per questo paziente si possono effettuare ulteriori considerazioni. La stima dei parametri k_0 , k_e , k e y_b^{CB} supera il test del t-value al 95% di confidenza, si osserva anche un valore relativamente basso di σ sul valore normalizzato per tali

parametri. Anche il test del χ^2 al 95% di confidenza viene superato (Tabella 4.8), ottenendo un buon fitting dei dati e quindi una rappresentazione adeguata del fenomeno.

Tabella 4.7: Stime parametri paziente M; confronto tra stime associate ai pool e stime con i dati singoli.

(a) Stime M.

	Stima individuale	Normalizzati	Stima pool	Norm. pool.	θ_0
k_0	0.00952	0.163	0.0093	0.159	0.0584
k_1	0.01997	0.569	0.03025	0.862	0.0351
k_2	0.02419	0.157	0.03348	0.217	0.154
k_3	2.84×10^{-9}	0.892	8.5×10^{-9}	2.673	3.18×10^{-9}
k_4	0	0	0	0	0.0236
k_5	0.00175	1.286	0.00132	0.973	0.00136
k_6	0.1441	1.624	0.1715	1.934	0.0887
k_e	0.02005	2.541	0.01884	2.388	0.00789
D	301.4	5.049	286.9	4.806	59.7
t_{max}	56.27	1.745	65.67	2.036	32.25
k	0.54	0.752	0.57	0.796	0.719
y_b^{CB}	2.51	0.45	2.63	0.472	5.58

(b) Tabella residui

	Costanti	Variabili	SWR
y_{AG}	10.11	12.09	7.578
y_{CB}	10.11	12.09	6.152
z_{UL}	6.433	-20.97	1.284
z_H	6.433	-20.97	1.459
z_L	6.433	-20.97	1.162
TOT	39.51	-38.74	17.64
Φ^{ML}			18.409

Le matrici di varianza-covarianza e la matrice di correlazione del paziente G sono riportate in Tabella 4.9. È presente forte correlazione tra k_1 con k_4 e k_6 e tra k_4 con k_6 i canali dei multimeri UL, k_2 con k_3 , i multimeri H e L, mentre forte anticorrelazione tra k_1 con k_2 e k_3 , i canali di produzione dei diversi multimeri, tra k_2 con k_4 e k_6 , tra k_3 con k_4 e k_6 ed infine la completa anticorrelazione tra i parametri del rilascio D e t_{max} . Tramite quest'ultima stima è confermata la criticità e la difficile caratterizzazione dei parametri della fase di proteolisi osservata nel capitolo precedente e conseguentemente i limiti sull'identificabilità del set parametrico del singolo paziente avendo a disposizione la coppia di dati di elettroforesi e di VWF:AG/CB.

In Figura 4.8 vengono riportati, titolo esemplificativo, i profili predetti dalla stima dei dati singoli di AG e CB, delle unità multimeriche e delle frazioni ponderali del paziente G.

I profili di antigene e collagene presentano un buon fitting sui dati sperimentali, ma rispetto ai dati medi presentano dei valori maggiori dei multimeri con un picco più alto del VWF:CB di cir-

Tabella 4.8: Stime parametri paziente G ; confronto tra stime associate ai pool e stime con i dati singoli. I valori del t -value al 95% di confidenza sono riportati tra parentesi tonde, quelli in grassetto superano il test.

(a) Stime G

	Stima individuale	Normalizzati	Stima pool	Norm. pool.	θ_0	σ
k_0	0.06282 (3.1)	1.076	0.0712	1.219	0.0584	0.17
k_1	0.2715 (0.049)	7.736	0.2067	5.888	0.0351	77
k_2	0.06986 (0.034)	0.454	0.03425	0.222	0.154	6.5
k_3	3.54×10^{-8} (1×10^{-8})	11.12	4.42×10^{-8}	13.9	3.18×10^{-9}	$5.3 \times 10^{+8}$
k_4	0.2673 (0.051)	11.32	0.3518	14.91	0.0236	110
k_5	0.00081 (0.61)	0.596	0.00116	0.851	0.00136	0.48
k_6	0.4364 (0.049)	4.92	0.4657	5.25	0.0887	49
k_e	0.01024 (7.4)	1.298	0.00979	1.241	0.00789	0.086
D	72.2 (3.2×10^{-7})	1.209	59.89	1.003	59.7	$1.9 \times 10^{+6}$
t_{\max}	33.08 (6.8×10^{-7})	1.026	31.12	0.965	32.25	$7.4 \times 10^{+5}$
k	0.62 (4.8)	0.86	0.64	0.885	0.719	0.088
y_b^{CB}	2.39 (5.4)	0.429	2.47	0.442	5.58	0.039
Ref. t-value	1.6958					

(b) Tabella residui

	Costanti	Variabili	SWR
y_{AG}	10.11	12.09	9.565
y_{CB}	10.11	12.09	6.829
z_{UL}	6.433	-20.97	1.948
z_H	6.433	-20.97	0.9674
z_L	6.433	-20.97	0.7065
TOT	39.51	-38.74	20.02
Φ^{ML}			20.789
χ^2	40.031	Ref. χ^2 95% conf	44.985

Tabella 4.9: Matrici di varianza-covarianza e di correlazione per la stima associata ai pool con tutti dati singoli del paziente G; in grassetto sono evidenziate le correlazioni/anticorrelazioni.

(a) Matrice varianza-covarianza												
	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e	D	t_{max}	k	y_b^{CB}
k_0	0.0287											
k_1	0.842	5890										
k_2	-0.152	-496	42.4									
k_3	-9.99×10^6	-4.06×10^{10}	3.42×10^9	2.82×10^{17}								
k_4	2.1	8310	-707	-5.71×10^{10}	1.18×10^4							
k_5	-0.000989	8.62	-0.811	-5.11×10^7	14.2	0.231						
k_6	0.975	3720	-313	-2.59×10^{10}	5240	4.64	2380					
D	-0.00929	-0.691	0.0873	4.98×10^6	-1.41	-0.0206	-0.462	0.00738				
k	1.91	4370	-416	-1.6×10^{10}	7350	9.49	3310	-0.972	3.45×10^{12}			
t_{max}	-0.812	-1780	170	6.55×10^9	-3000	-3.88	-1350	0.407	-1.41×10^{12}	5.5×10^{11}		
k	-0.000744	1.61	-0.136	-1.09×10^7	2.31	0.0144	1.02	-0.00017	1.46	-0.531	0.00774	
y_b^{CB}	-0.00053	0.124	-0.0123	-5.93×10^5	0.221	0.00617	0.0532	-3.85×10^{-5}	0.112	-0.0408	0.00302	0.0015

(b) Matrice correlazione												
	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e	D	t_{max}	k	y_b^{CB}
k_0	1											
k_1	0.0647	1										
k_2	-0.138	-0.994	1									
k_3	-0.111	-0.996	0.999	1								
k_4	0.114	0.995	-0.999	-0.989	1							
k_5	-0.0121	0.234	-0.259	-0.2	0.271	1						
k_6	0.118	0.994	-0.987	-0.999	0.987	0.198	1					
k_e	-0.638	-0.105	0.156	0.109	-0.15	-0.498	-0.11	1				
D	6.08×10^{-6}	3.07×10^{-5}	-3.45×10^{-5}	-1.63×10^{-5}	3.64×10^{-5}	1.06×10^{-5}	3.66×10^{-5}	-6.09×10^{-6}	1			
t_{max}	-6.46×10^{-6}	-3.14×10^{-5}	3.52×10^{-5}	1.66×10^{-5}	-3.72×10^{-5}	-1.09×10^{-5}	-3.74×10^{-5}	6.39×10^{-5}	-1	1		
k	-0.0499	0.238	-0.238	-0.233	0.241	0.341	0.237	-0.0225	8.91×10^{-6}	-8.14×10^{-6}	1	
y_b^{CB}	-0.0808	0.0418	-0.049	-0.0289	0.0525	0.332	0.0282	-0.0116	1.57×10^{-6}	-1.42×10^{-6}	0.886	1

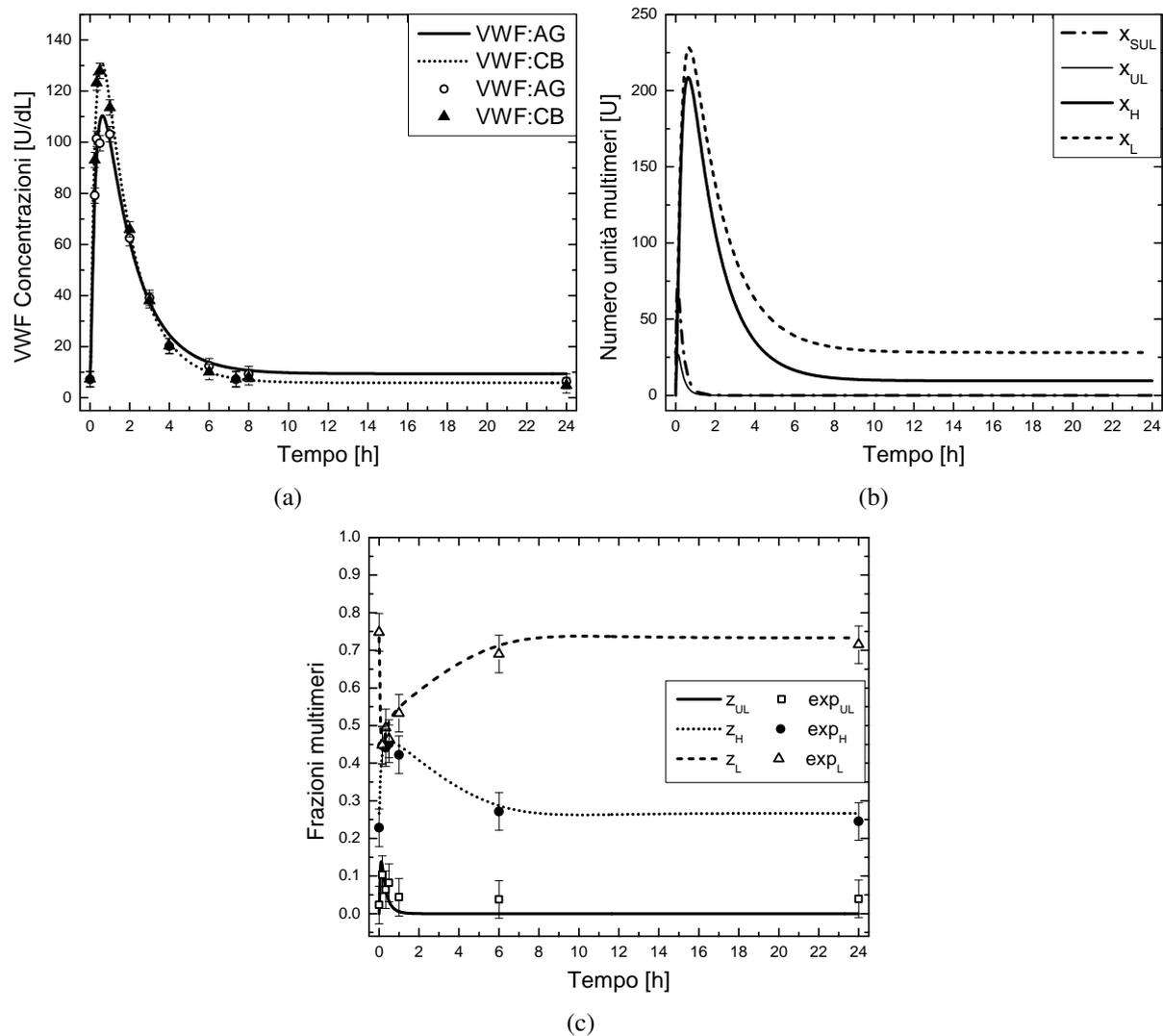


Figura 4.8: Profili di VWF:AG e VWF:CB predetti dal modello per il paziente G (a); profili delle unità multimeriche predette dal modello per il paziente G (b) e fitting delle frazioni di gel elettroforesi (c).

ca 30 U/dL rispetto al VWF:AG; per quanto riguarda i profili delle frazioni (Figura 4.8c), anche in questo caso si ha un buon fitting seppur venga predetto un picco più elevato dei multimeri UL, la quale è una conseguenza dei valori maggior di VWF:CB (Figura 4.8b).

4.6 Analisi di sensitività e di informazione sulle frazioni

Nelle Figure 4.9 e 4.10 sono riportati i profili delle matrici di sensitività delle frazioni multimeriche z_{UL} , z_H e z_L rispettivamente per la VWD 2B e Vicenza².

Entrambe le tipologie esibiscono per i multimeri UL delle curve aventi una dinamica molto ve-

²L'analisi di sensitività è stata condotta tramite uno scaling di \mathbf{Q}_r utilizzando un valore molto piccolo $\varepsilon = 10^{-5}$, per risolvere dei problemi numerici rispetto alle divisioni per quantità molto piccole nel calcolo delle sensitività delle risposte delle frazioni.

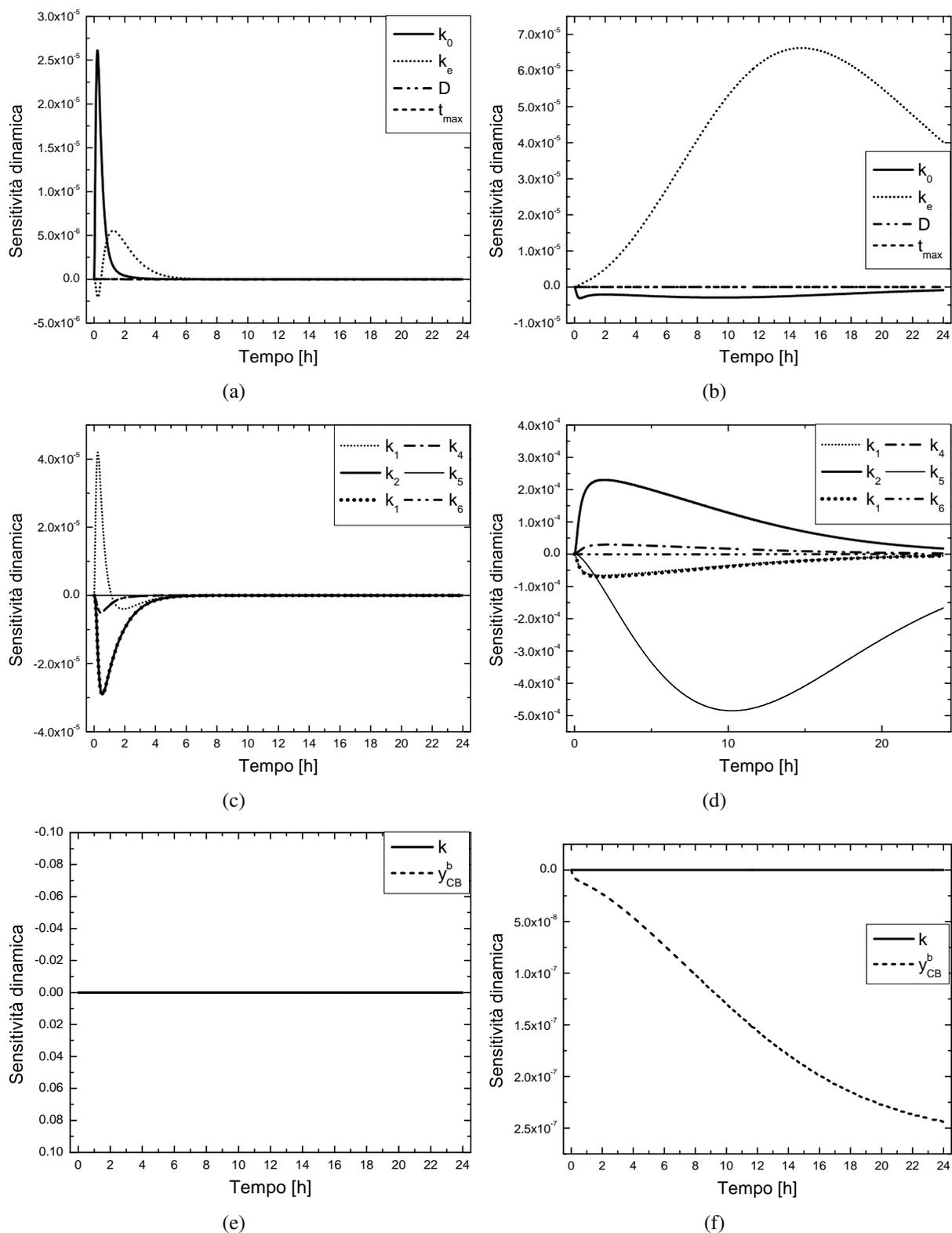


Figura 4.9: Analisi di sensitività per le risposte di frazioni dei multimeri predette dal modello 6M per il pool VWD 2B. I grafici a sinistra (a, c, e) sono riferiti ai multimeri UL, mentre a destra (b, d, f) ai multimeri H (le quali sono simmetriche rispetto alle ascisse per i multimeri L)

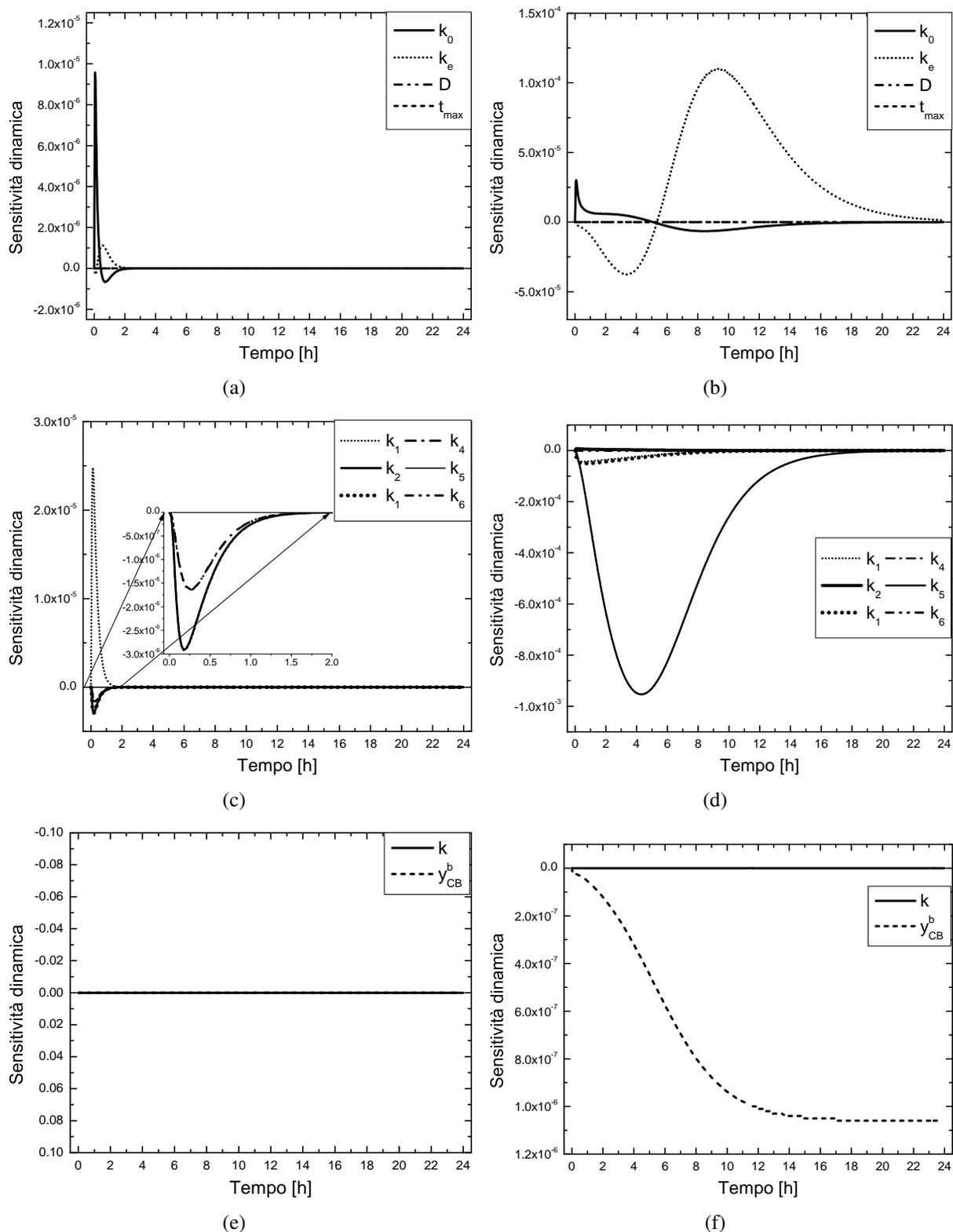


Figura 4.10: Analisi di sensitività per le risposte di frazioni dei multimeri predette dal modello 6M per il pool VWD Vicenza. I grafici a sinistra (a, c, e) sono riferiti ai multimeri UL, mentre a destra (b, d, f) ai multimeri H (le quali sono simmetriche rispetto alle ascisse per i multimeri L).

loce, con dei picchi nei primi istanti temporali fino alle 2-3 h, sia per i parametri di rilascio ed eliminazione (Figure 4.9a, 4.10a) che per la fase di proteolisi, con valori maggiori per i parametri k_1 , k_2 e k_3 (Figure 4.9c, 4.10c).

Gli andamenti sono diversi invece per il multimeri H ed L: poiché le frazioni z_H e z_L hanno valori dominanti rispetto alle z_{UL} , gli andamenti dei profili di sensitività sono tra di loro simmetriche rispetto all'origine, quindi vengono riportati solamente i profili rispetto alle risposte z_H . La dipendenza temporale di quest'ultime verso i parametri di rilascio sono molto piccole per D e t_{max} , e rispecchiano delle dinamiche più veloci per k_0 ; per quanto riguarda l'eliminazione le dinamiche sono più marcate e presentano dei picchi caratteristici verso le ore 14 per la VWD 2B (Figura 4.9b) e con delle oscillazioni alle 3 e 9 h per VWD Vicenza (Figura 4.10b). Osservando gli andamenti dei parametri della proteolisi si riscontrano le stesse caratteristiche viste per le sensitività di VWF:AG e VWF:CB nel §3.5, soprattutto il parametro k_5 , il quale rappresenta il canale tra i multimeri H ed L, avente una dinamica che è tipica delle risposte delle tipologie prese in considerazione (Figure 4.9d ed 4.10d).

Per entrambe le tipologie i parametri della correzione hanno sensitività nulla rispetto ai multimeri UL (Figura 4.9e, 4.10e), ma per i multimeri H ed L è presente una piccola dipendenza per y_b^{CB} (Figure 4.9f ed 4.10f).

I profili delle tracce delle matrici di informazione rispetto alle risposte delle frazioni ponderali sono riportati in Figura 4.11. Gli andamenti rispecchiano i profili delle sensitività dei multimeri H ed L, poiché tali frazioni sono dominanti durante l'esperimento, e presentano dinamiche

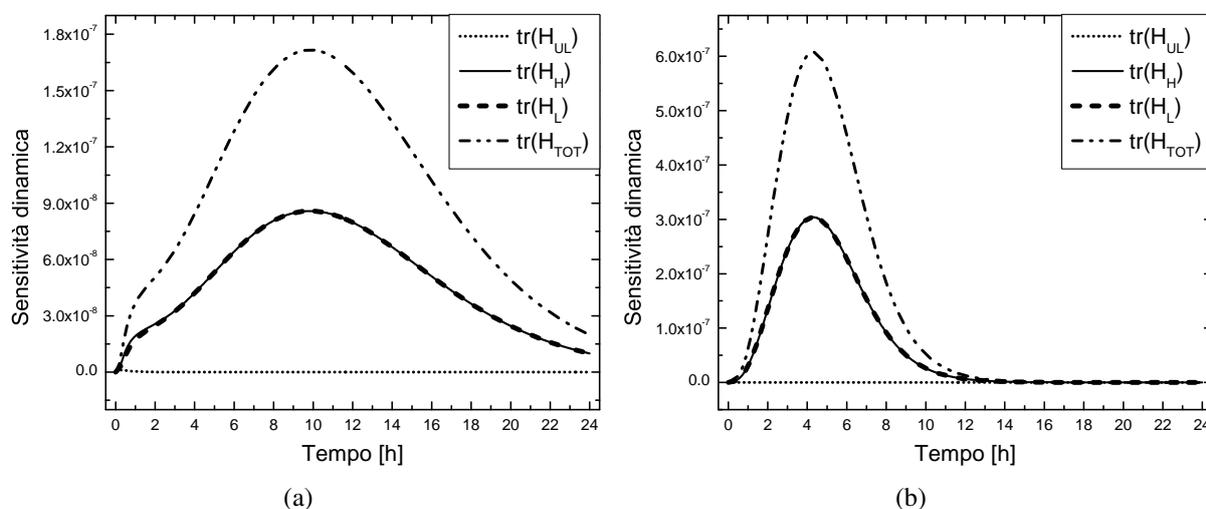


Figura 4.11: Profili delle tracce delle matrici di informazioni rispetto alle risposte delle frazioni di multimeri, per i pazienti VWD 2B (a) e per i pazienti VWD Vicenza (b).

caratterizzanti la tipologia della malattia, ad esempio si noti il valore massimo dei 2B verso le 10 h (Figura 4.11a), mentre per i Vicenza il massimo è spostato verso le 4 h (Figura 4.11b). Un aspetto importante da sottolineare è che i profili delle sottomatrici di proteolisi (vedi §3.6) hanno lo stesso andamento delle matrici totali, si può affermare quindi con certezza che il tipo di informazione fornita dalle frazioni dei diversi multimeri è specifica per la fase di proteolisi. Dalle analisi di sensitività e di informazione precedenti si possono trarre le seguenti conclusioni: l'elettroforesi, di conseguenza i dati delle frazioni, è uno strumento di supporto alla diagnosi perché fornisce molta informazione sulla qualità della distribuzione dei multimeri, per gli istanti iniziali, aiutando in fase di stima ad identificare i flussi nella fase di proteolisi, anche se non sempre tali stime associate ai dati di VW:AG e VW:CB forniscono una diagnosi sicura.

4.7 Riparametrizzazione

Le frazioni multimeriche, come si è visto, forniscono un'informazione aggiuntiva, ma il modello 6M presenta ancora i seguenti limiti:

- risulta essere presente una condizione di sovrapparametrizzazione;
- il sistema di equazioni non è identificabile ed i parametri assumono in fase di stima dei valori non univoci;
- non avendo delle statistiche in uscita dallo stimatore non si ha a disposizione una misura sulla precisione delle stime parametriche.

Per risolvere queste problematiche si è effettuata una riparametrizzazione in fase di stima del modello (§2.5), basata sull'osservazione del significato fisiologico di ciascun parametro rispetto alle diverse fasi, in riferimento agli andamenti delle stime precedenti.

La riparametrizzazione è stata condotta mantenendo fissati i parametri D e t_{max} del rilascio, k e y_b^{CB} della correzione, secondo i valori delle stime preliminari del capitolo precedente. I valori fissati utilizzati per ogni tipologia di dati sono riassunti in Tabella 4.10; il valore inferiore di D per i pazienti conferma il riconoscimento del modello di un deficit nella secrezione del fattore VWF nella malattia.

Per quanto riguarda i parametri rimanenti le stime vengono effettuate sul set di parametri:

$$\boldsymbol{\theta}' = [\theta'_1, k_1, k_2, k_3, k_4, k_5, k_6, \theta'_2] \quad (4.3)$$

Tabella 4.10: Parametri fissati utilizzati nelle stime riparametrizzate.

	0	Non-0	VWD Vicenza	VWD 2B	VWD 2A
D	100	100	40	40	40
t_{max}	30	30	30	30	30
k	1.04	0.96	0.719	0.229	0.044
y_b^{CB}	48.8	66.54	5.58	10.05	3.93

Tabella 4.11: Stime parametriche effettuate dopo la riparametrizzazione.

	0	Non-0	VWD Vicenza	VWD 2B	VWD 2A
k₀	0.0480	0.0564	0.075	0.083	0.091
k₁	0.0108	9.96×10^{-9}	0.018	0.0166	0.0068
k₂	0.0537	0.0535	0.133	0.00513	0.0012
k₃	0.0041	0.0184	1.19×10^{-9}	0.0014	0.0177
k₄	0.001255	4.71×10^{-4}	0.0269	0.0024	0.0032
k₅	7.17×10^{-5}	1.43×10^{-4}	0.00138	7.43×10^{-5}	3.9×10^{-4}
k₆	0.0237	0.1297	0.1467	0.1613	0.0150
k_e	1.506×10^{-3}	6.45×10^{-4}	0.00757	0.00286	0.00109

dove i parametri θ'_1 e θ'_2 sono stati riparametrizzati nel modo seguente:

$$\begin{aligned}\theta'_1 &= \frac{k_0 \cdot D}{k_e} \\ \theta'_2 &= k_e \cdot D\end{aligned}\tag{4.4}$$

Conducendo le stime adottando tale riparametrizzazione il numero di parametri da stimare diminuisce (da 12 a 8) e i profili delle matrici di sensitività cambiano (eq. 2.26) senza alterare la struttura del modello. Per osservare i benefici della riparametrizzazione si sono ripetute le stime per ogni tipologia di dati a disposizione, utilizzando anche i dati medi dei soggetti sani 0 e non-0 oltre ai dati di pazienti affetti da VWD.

Tutte le stime, i valori nominali, gli intervalli di stima, le relative statistiche ottenute, le matrici di correlazione e di varianza-covarianza sono riportate in Appendice B nelle rispettive Tabelle B.1-B.5 per ogni set di dati sperimentali, mentre in Tabella 4.11 vengono riportate le sintesi delle stime delle cinque classi di dati dopo la riparamerizzazione.

Se si osserva il test sul χ^2 , nella Tabella di ogni tipologia, si nota come il fitting sia ottimale per i soggetti 0, VWD Vicenza e 2B, e non statisticamente soddisfacente per i soggetti non-0 e 2A; per quanto riguarda i residui e i valori della funzione obiettivo essi presentano valori simili alle stime preliminari di Tabella 3.1. Rispetto alle stime preliminari ottenute dal sistema non riparametrizzato queste nuove stime presentano degli scostamenti nei valori dei parametri. Si

possono rilevare dei cambiamenti:

- per i soggetti 0 il parametro k_2 aumenta del 30%;
- per i non-0 si hanno cambiamenti di circa il $\pm 40\%$ dei parametri, tranne che per k_5 e k_e ;
- per la VWD Vicenza si ha una stima inferiore del 30% di k_1 , k_2 e k_3 ;
- per i VWD 2B circa il $\pm 20\%$ dei parametri k_3 e k_4 ;
- per i 2A non si hanno variazioni significative nelle stime.

L'ottenimento delle statistiche significa che il sistema così riparametrizzato è identificabile. Il test del t-value al 95% di confidenza è superato per tutti i casi dalle stime dei parametri del rilascio k_0 e di eliminazione k_e , quindi la riparametrizzazione ha avuto dei benefici per questi due parametri anche osservando i bassi valori di deviazione standard.

Per i parametri della proteolisi invece si hanno dei valori di varianza-covarianza molto elevati ($10^5 - 10^8$) il che è sicuramente legato all'elevata correlazione o anticorrelazione presente tra questi parametri. Analizzando le matrici di correlazione, riportate nel dettaglio nelle Tabelle B.1-B.5, si possono osservare le seguenti correlazioni tra le coppie:

- per soggetti non-0, VWD 2B e VWD 2A: k_1-k_4 ;
- VWD Vicenza: k_1-k_4 e k_2-k_6 ;
- i soggetti 0 non presentano correlazioni;

mentre per quanto riguardano le anticorrelazioni tutte le tipologie presentano anticorrelazione tra k_0 e k_e e tra k_1 e k_2 , e nello specifico tra le coppie:

- soggetti 0: k_4-k_6
- soggetti non-0: k_1-k_6 , k_2-k_4 e k_4-k_6 ;
- VWD Vicenza: k_1-k_6 e k_2-k_4 ;
- VWD 2B: k_1-k_6 e k_4-k_6 ;
- VWD 2A: k_1-k_6 , k_2-k_4 e k_4-k_6 .

Dunque nella maggior parte dei casi i canali delle proteolisi problematici risultano essere quelli che implicano: i flussi di produzione ed eliminazione dei multimeri UL rispetto ai SUL e H, i quali, ricordando i profili delle frazioni, sono presenti in quantità minori e rispetto ai restanti hanno vita breve, avendo una clearance inferiore alle 2 h; i flussi di produzione e consumo dei multimeri H rispetto ai multimeri SUL e UL. Per quanto riguardano i flussi interessanti i multimeri L non sono presenti correlazioni data la natura di identificabilità delle misure di VWF:AG e CB dimostrata nel §3.6.

4.8 Simulazione stocastica

Successivamente è stato condotto un breve studio di sensitività per valutare la variabilità delle risposte misurate di VWF:AG e CB e z_i , al variare del valore dei parametri di proteolisi all'interno di una variabilità prefissata. È stata condotta una simulazione stocastica con un numero di 100 scenari, variando i parametri da k_1 a k_6 secondo una distribuzione random uniforme pari ad una percentuale di ciascun valore dei parametri θ_i presenti in Tabella 4.11.

Si osserva che, mediamente per le diverse tipologie, variando tale percentuale di variazione dei parametri, le risposte delle frazioni multimeriche assumono degli andamenti significativi dal punto di vista fisiologico, cioè con dinamiche prive di oscillazioni, per una percentuale inferiore del $\pm 40\%$, mentre le risposte di antigene e collagene rimangono all'interno dell'intervallo di SE di ogni pool anche per percentuali di variazione maggiori ($\pm 80-90\%$).

Vengono riportati i profili delle risposte predette per le simulazioni stocastiche, con una variazione del $\pm 40\%$ sui parametri di proteolisi, dei soggetti normali 0 in Figura 4.12 e VWD Vicenza in Figura 4.13. Si nota come accennato precedentemente, che per tutte le tipologie i profili di antigene e collagene rimangono all'interno di una variabilità molto stretta (± 5 U/dL), eccetto per il collagene della VWD 2B che presenta una banda più larga, circa $\pm 10-20$ U/dL (Figura 4.12a,b, Figura 4.13a,b), rispetto alla variabilità molto elevata delle risposte predette delle frazioni multimeriche, soprattutto sui multimeri H ed L che arriva fino a ± 0.2 in termini di frazioni, cioè una variazione del $\pm 20\%$ sulle frazioni totali (Figura 4.12c,d,e, Figura 4.13c,d,e), che corrisponde ad un intervallo di circa $\pm 30-50$ U/dL in termini di unità multimeriche, il quale è più ampio dell'intervallo di variabilità delle diverse tipologie di VWD.

Questi risultati mettono in evidenza il fatto che, avendo a disposizione entrambe le modalità di misura VWF:AG/CB e z_i , dei sampling tramite delle misure precise di antigene e collagene possono portare a degli errori di stima del set parametrico più elevati rispetto a delle misure precise di frazioni ponderali; il che significa, da un punto di vista contrario, che le misure di frazioni multimeriche portano ad un'identificazione del set parametrico con una maggiore precisione, questo spiega il perché le analisi di gel-elettroforesi vengono utilizzate dal personale medico per acquisire maggiori informazioni sul tipo di VWD sofferta dal paziente.

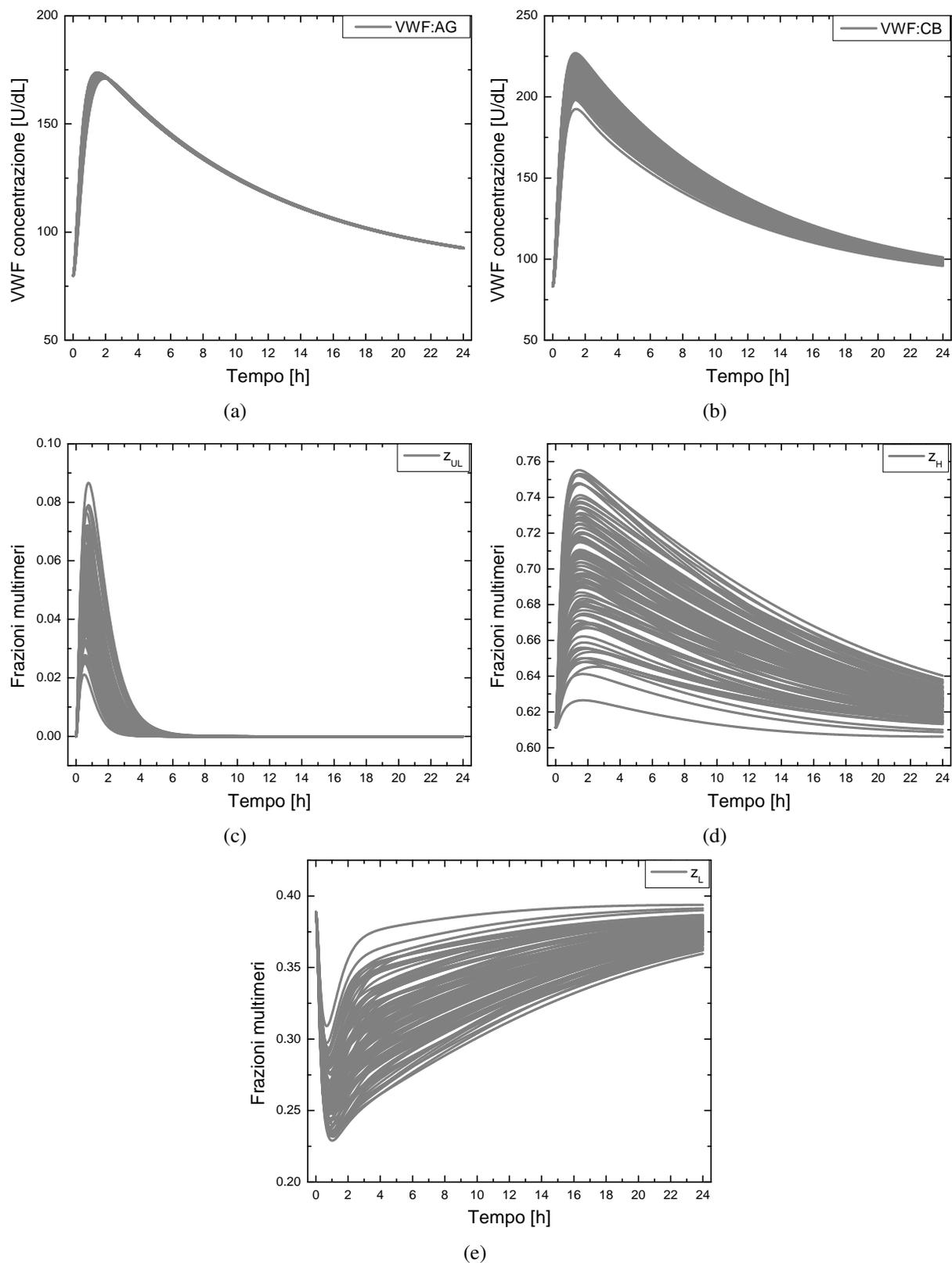


Figura 4.12: Profili predetti dal modello tramite una simulazione stocastica variando del $\pm 40\%$ i parametri della proteolisi rispetto ai soggetti normali. Nelle sottofigure sono riportati in ordine: (a) VWF:AG, (b) VWF:CB, (c) frazioni multimeri z_{UL} , (d) frazioni multimeri z_H , (e) frazioni multimeri z_L .

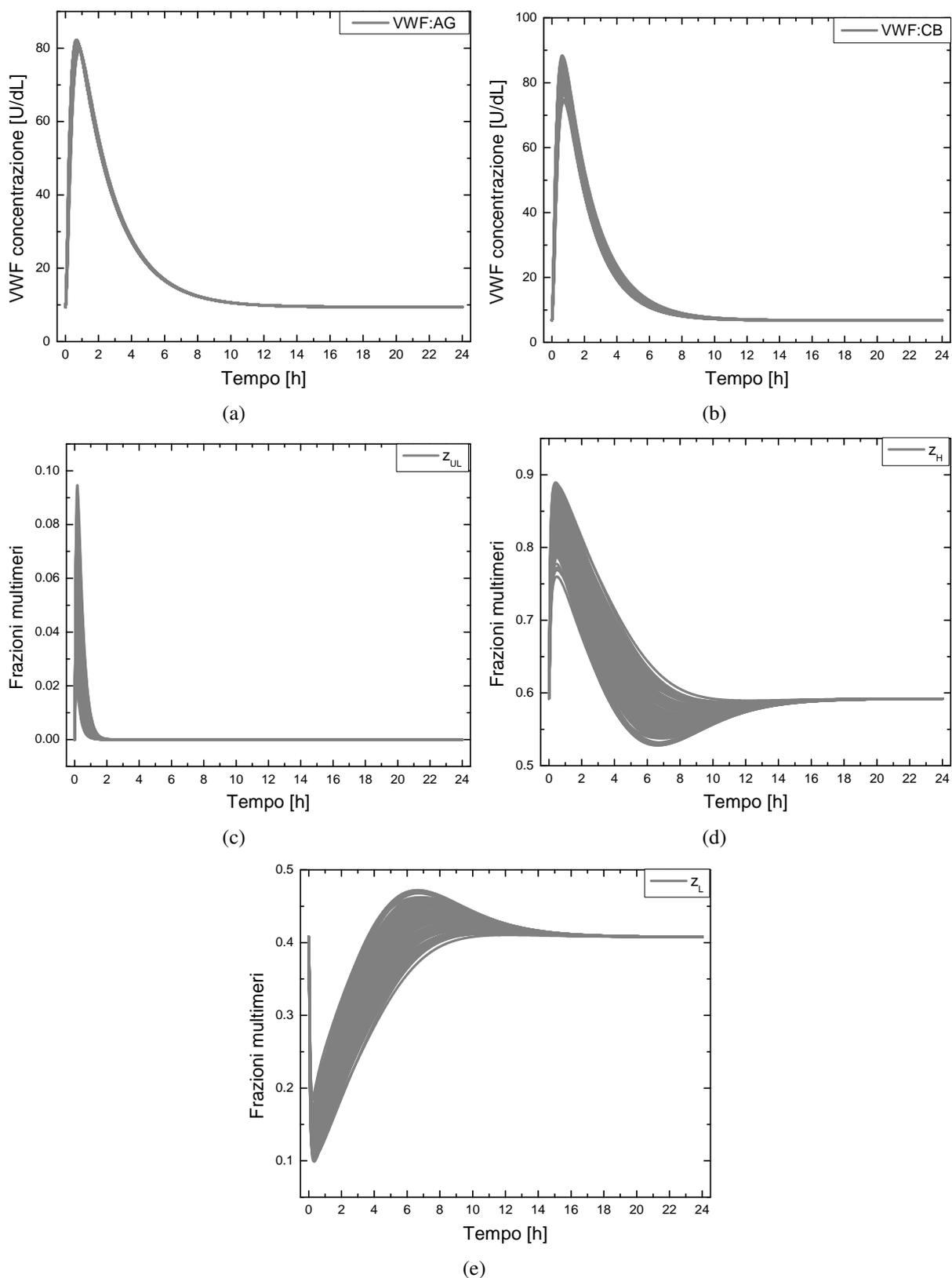


Figura 4.13: Profili predetti dal modello tramite una simulazione stocastica variando del $\pm 40\%$ i parametri della proteolisi rispetto ai pazienti VWD Vicenza. Nelle sottofigure sono riportati in ordine: (a) VWF:AG, (b) VWF:CB, (c) frazioni multimeri z_{UL} , (d) frazioni multimeri z_H , (e) frazioni multimeri z_L .

4.9 Conclusioni

Il modello 6M riconosce molto bene gli andamenti caratteristici di VWF:AG e CB di ogni tipologia di malattia, ma se si vogliono ottenere delle stime individuali più precise del singolo paziente è necessario avere a disposizione anche le frazioni multimeriche del paziente, come nel caso del paziente G. di §4.4. Purtroppo il procedimento sperimentale per ottenere le lastre di elettroforesi ed i tempi di elaborazione sono molto onerosi (per avere dei risultati si devono attendere anche 10 giorni).

Le misure di antigene e collagene sono molto più economiche e si hanno a disposizione i dati a distanza di pochi minuti dal campionamento. Dal punto di vista diagnostico sarà importante ottenere delle valutazioni rapide ed economiche, quindi, al fine di utilizzare il modello per la diagnosi della VWD sarà necessario utilizzare i soli dati di VWF:AG e CB.

Per aumentare la precisione delle stime parametriche del singolo paziente senza voler compiere l'analisi di elettroforesi, sarà utile in futuro effettuare un MBD_{oE} per adottare un'ulteriore riparametrizzazione del modello che riduca la forte correlazione dei flussi citati nel §4.6, ripercorrendo l'iter di progettazione degli esperimenti presentato nel Capitolo 2 in Figura 2.1, in modo tale da ottenere delle predizioni precise delle frazioni solamente tramite dati di antigene e collagene.

Conclusioni

L'obiettivo principale di questo lavoro di Tesi è quello di identificare, attraverso un approccio ingegneristico, un modello meccanicistico per lo studio della VWD. L'importanza di avere a disposizione un modello matematico affidabile che descriva in modo quantitativo la distribuzione dei multimeri di VWF nel tempo, rappresenta un notevole supporto all'attività medica per aumentare la comprensione dei complessi fenomeni fisiologici che avvengono nei pazienti affetti da VWD, per fornire gli strumenti necessari in modo tale da proporre un protocollo per una diagnosi più sicura e veloce e, non per ultimo, per poter consentire delle simulazioni in silico in funzione di una terapia più efficiente.

Le tecniche utilizzate in questa Tesi forniscono un prezioso aiuto all'identificazione del modello, consentendo la massimizzazione del contenuto informativo degli studi clinici, con lo scopo di discriminare le diverse strutture proposte e per effettuare la stima dei parametri del modello con la massima precisione. Inoltre risultano particolarmente utili per progettare le condizioni sperimentali con il massimo contenuto informativo senza gravare sulla salute del paziente affetto da VWD.

I modelli proposti 6M e 7 predicano nel modo ottimale i profili di VWF:AG e VWF:CB, per di più il modello 6M riesce a descrivere con maggior dettaglio le singole specie multimeriche di ogni variante della malattia. Si è osservato come entrambi i modelli abbiano una condizione di sovrapparametrizzazione che determina un problema di non identificabilità parametrica. Il set parametrico, in fase di stima, non converge verso una soluzione univoca, soprattutto rispetto ai parametri della fase di proteolisi. Si è visto come una riparametrizzazione risolva il problema di non identificabilità pur rimanendo un'elevata correlazione tra i parametri della fase proteolitica corrispondenti al canale dei multimeri ad alto peso molecolare.

Il risultato ottenuto da questo lavoro è quello di ottenere, tramite la procedura di diagnosi mediante modello descritta nel Capitolo 4, delle diagnosi affidabili in un certo numero di casi di pazienti. Il metodo richiede il supporto di dati individuali di elettroforesi, interpretati attraverso tecniche di analisi d'immagine. Tuttavia, visto che l'elettroforesi è una misura complessa e dispendiosa per i tempi che richiede, uno degli obiettivi della ricerca futura sarà quello di fare

principalmente affidamento sulle misure di antigene e collagene, utilizzando il modello 7 che non necessita di dati relativi i diversi canali multimerici. In quest'ottica potrebbe rivelarsi di grande efficacia l'uso combinato di tecniche di riparametrizzazione e progettazione degli esperimenti.

Un'ulteriore sviluppo futuro riguarda l'estensione del metodo di diagnosi alla tipologia VWD 1, della quale non si hanno a disposizione dati sperimentali.

Nomenclatura

Simboli generici

A	=	area di densità ottica [U]
AUC	=	<i>area under curve</i>
c_{ij}	=	correlazione tra la colonna i -esima e j -esima di \mathbf{P}_E
F_i	=	flusso i -esimo [U]
N_{exp}	=	numero di esperimenti
N_M	=	numero di risposte stimate dal modello
N_m	=	numero di modelli
n_{sp}	=	numero di campionamenti
N_U	=	numero delle variabili manipolabili dinamiche
N_W	=	numero delle variabili manipolabili costanti
N_y	=	numero delle uscite del modello
N_θ	=	numero di parametri del modello
p_{Ei}	=	elemento dell' i -esima colonna di \mathbf{P}_E
q_{ij}	=	elemento della matrice di sensitività dinamica locale
s_{ij}	=	ij -esimo elemento dell'inverso della matrice Σ
t	=	tempo [h]
t_i	=	t -value dell' i -esimo esperimento
V_d	=	Volume di distribuzione [L]
v_{ij}	=	covarianza del parametro i -esimo rispetto al parametro j -esimo
x^i	=	multimeri i -esimi [U]
x_b^i	=	valore basale degli i -esimi multimeri [U]
y	=	generica risposta del modello
y^i	=	risposta i -esima del modello per la VWD [U·dL ⁻¹]
y_b^i	=	concentrazione basale dell' i -esima risposta [U·dL ⁻¹]
z_i	=	frazione ponderale i -esima

Simboli greci

α	=	livello statistico di confidenza
β	=	termine costante della funzione obiettivo
γ	=	parametro di varianza eteroschedastica
ε_y	=	valore arbitrariamente piccolo sulle risposte
ε_θ	=	valore arbitrariamente piccolo sui parametri
η_i	=	valor medio degli elementi dell' i -esima colonna di \mathbf{P}_E
θ_i	=	parametro i -esimo
κ	=	intervallo di confidenza dell' i -esimo parametro
μ_i	=	fattore di normalizzazione del parametro i -esimo
ν	=	gradi di libertà
σ	=	deviazione standard
σ_y	=	deviazione standard sulla risposta y
σ_θ	=	deviazione standard sul parametro θ
$\tilde{\sigma}_{r,s}$	=	elemento r,s -esimo dell'inverso della matrice di varianza-covarianza Σ_y
τ	=	durata dell'esperimento
Φ	=	funzione obiettivo
Φ^I	=	funzione obiettivo del test di identificabilità <i>a-posteriori</i>
Φ^{PE}	=	funzione obiettivo della stima parametrica
Φ^{ML}	=	funzione obiettivo di massima verosimiglianza
Φ_{TOT}	=	funzione obiettivo totale
χ^2	=	<i>lack-of fit test</i>
χ_{RIF}	=	valore di riferimento sul test <i>lack-of fit</i>
$\psi(\cdot)$	=	generica metrica relativa ad una matrice (\cdot)
ω	=	parametro del modello di varianza

Matrici e vettori

\mathbf{A}	=	matrice di riparametrizzazione [$N_{\theta'} \times N_{\theta'}$]
\mathbf{C}	=	matrice di correlazione dei parametri [$N_\theta \times N_\theta$]
$\mathbf{H}(\theta)$	=	matrice di informazione dinamica [$N_\theta \times N_\theta$]
$\mathbf{H}(\theta, \varphi)$	=	matrice di informazione dinamica attesa dall'esperimento [$N_\theta \times N_\theta$]
\mathbf{H}_θ^0	=	matrice di informazione iniziale [$N_\theta \times N_\theta$]

$\mathbf{I}(\boldsymbol{\theta})$	=	matrice di informazione di Fisher $[N_{\theta} \times N_{\theta}]$
$\mathbf{M}(\boldsymbol{\theta})$	=	modello nello spazio degli stati (<i>state space model</i>) $[\cdot]$
\mathbf{P}_E	=	matrice di stimabilità $[n_{sp}N_y \times N_{\theta}]$
$\mathbf{Q}(t)$	=	matrice di sensitività dinamica locale $[N_y \times N_{\theta}]$
\mathbf{Q}_r	=	matrice di sensitività dinamica rispetto alla risposta r -esima $[n_{sp} \times N_{\theta}]$
\mathbf{r}	=	matrice dei residui assoluti $[N_y \times n_{sp}]$
$\mathbf{S}_y(t_k)$	=	matrice di sensitività locale valutata al tempo di campionamento t_k $[N_y \times N_{\theta}]$
\mathbf{t}_{sp}	=	vettore dei tempi di campionamento $[N_{sp} \times 1]$
$\mathbf{u}(t)$	=	vettore delle variabili manipolate $[N_U \times 1]$
\mathbf{V}_{θ}	=	matrice di varianza-covarianza sui parametri $[N_{\theta} \times N_{\theta}]$
\mathbf{w}	=	vettore delle variabili manipolate costanti nel tempo $[N_W \times 1]$
\mathbf{W}	=	matrice peso generica $[\cdot]$
\mathbf{W}_{θ}	=	matrice peso sui parametri $[N_{\theta} \times N_{\theta}]$
\mathbf{W}_y	=	matrice peso sui parametri $[N_y \times N_y]$
$\mathbf{x}(t)$	=	vettore delle variabili di stato $[N_x \times 1]$
$\dot{\mathbf{x}}(t)$	=	vettore delle derivate delle variabili di stato $[N_x \times 1]$
\mathbf{x}_0	=	vettore delle condizioni iniziali $[N_x \times 1]$
$\mathbf{y}(t)$	=	vettore delle risposte del modello $[N_y \times 1]$
$\hat{\mathbf{y}}(t)$	=	vettore delle risposte stimate del modello $[N_y \times 1]$
$\boldsymbol{\Theta}$	=	vettore dei parametri normalizzati del modello $[N_{\theta} \times 1]$
$\boldsymbol{\theta}$	=	vettore dei parametri del modello $[N_{\theta} \times 1]$
$\boldsymbol{\theta}_0$	=	vettore della stima iniziale/preliminare dei parametri del modello $[N_{\theta} \times 1]$
$\boldsymbol{\theta}'$	=	vettore dei parametri del modello ottenuto con la riparametrizzazione $[N_{\theta} \times 1]$
$\hat{\boldsymbol{\theta}}$	=	vettore delle stime parametriche ottenute sui parametri del modello $[N_{\theta} \times 1]$
$\boldsymbol{\Sigma}_i$	=	matrice di varianza-covarianza degli errori delle misure nell' i -esimo esperimento $[N_y \times N_y]$
$\boldsymbol{\Sigma}_y$	=	matrice di varianza-covarianza sulle misure sperimentali $[N_{\theta} \times N_{\theta}]$
$\boldsymbol{\Sigma}_{\theta}$	=	matrice di varianza-covarianza sui parametri $[N_{\theta} \times N_{\theta}]$
$\boldsymbol{\varphi}$	=	vettore dell'esperimento $[\cdot]$

Acronimi

ADP	=	adenosina difosfato
DAE	=	equazioni algebrico-differenziale (<i>differential algebraic equation</i>)
DDAVP	=	desmovasopressina
DoE	=	progettazione degli esperimenti (<i>design of experiments</i>)
ECM	=	matrice extracellulare (<i>extracellular matrix</i>)
H	=	multimeri a medio peso molecolate (<i>high molecular weight multimers</i>)
L	=	multimeri a basso peso molecolate (<i>low molecular weight multimers</i>)
MBD _{oE}	=	progettazione degli esperimenti basata su modello (<i>model based design of experiments</i>)
MIMO	=	sistemi <i>multiple-input multiple-output</i>
ML	=	massima verosimiglianza (<i>maximum likelihood</i>)
NHLBI	=	National Heart, Lung and Blood Institute
PE	=	stima dei parametri (<i>parameter estimation</i>)
PSE	=	<i>process system engineering</i>
SE	=	errore standard (<i>standard error</i>)
SGI	=	test di identificabilità strutturale globale (<i>structurally global identifiable</i>)
SLI	=	test di identificabilità strutturale locale (<i>structurally local identifiable</i>)
SUL	=	macromolecole multimeriche (<i>super ultra large multimers</i>)
TF	=	fattore tissutale (<i>tissutal factor</i>)
TTP	=	porpora trombotica trombocitopenica
TXA ₂	=	trombossano
UL	=	multimeri ad alto peso molecolate (<i>ultra large multimers</i>)
VWD	=	malattia di von Willebrand (<i>von Willebrand disease</i>)
VWF	=	fattore di von Willebrand (<i>von Willebrand factor</i>)
VWF:AG	=	antigene del fattore di von Willebrand binding
VWF:CB	=	collagene del fattore di von Willebrand
VWF:RC _o	=	ristocetina co-fattore del fattore di von Willebrand
SWR	=	somma dei residui pesati (<i>sum of weighted residuals</i>)

Appendice A

Stime parametriche ottenute per la diagnosi della VWD

In questa Appendice vengono riportate le stime parametriche, i valori nominali, le stime preliminari di riferimento e i limiti degli intervalli di stima delle stime effettuate per la diagnosi secondo la procedura del Capitolo 4 in Figura 4.1.

Le tabelle che seguono contengono le tre stime parametriche effettuate per ogni paziente, rispettivamente:

- Tabella A.1 paziente A;
- Tabella A.2 paziente B;
- Tabella A.3 paziente G;
- Tabella A.4 paziente M.

Ciascuna sottotabella rappresenta la stima con il pool:

- (a) VWD Vicenza;
- (b) VWD 2B;
- (c) VWD 2A.

Si ricorda che i fattori di normalizzazione adottati per ogni pool di dati θ_0 , sono le stime preliminari effettuate nel Capitolo 3.

Tabella A.1: Stime parametri in fase di diagnosi ottenute per il paziente A.

(a) VWD Vicenza

	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.03995	0.684	0.0584	1	0	100
k_1	0.1275	3.632	0.0351	1	0	100
k_2	0.1622	1.053	0.154	1	0	100
k_3	2.43×10^{-8}	7.642	3.18×10^{-9}	1	0	100
k_4	0.0573	2.428	0.0236	1	0	100
k_5	0.000364	0.268	0.00136	1	0	100
k_6	0.7846	8.846	0.0887	1	0	100
k_e	0.00941	1.192	0.00789	1	0	100
D	106	1.775	59.7	1	0	100
t_{\max}	39.04	1.211	32.25	1	0	100
k	0.74	1.031	0.719	1	0.6	1.5
y_b^{CB}	4.29	0.768	5.58	1	0.2	1.8

(b) VWD 2B

	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.0725	0.929	0.078	1	0	100
k_1	0.00984	0.589	0.0167	1	0	100
k_2	0.01206	2.369	0.00509	1	0	100
k_3	0.00245	1.065	0.0023	1	0	100
k_4	0	0	0.00439	1	0	100
k_5	0.000209	3.777	5.54×10^{-5}	1	0	100
k_6	0.1197	0.798	0.15	1	0	100
k_e	0.00302	1.078	0.0028	1	0	100
D	63.85	1.419	44.98	1	0	100
t_{\max}	28.1	0.925	30.39	1	0	100
k	0.22	0.979	0.229	1	0.8	2.2
y_b^{CB}	18.09	1.8	10.05	1	0.2	1.8

(c) VWD 2A

	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.5174	5.307	0.0975	1	0	100
k_1	0.000193	0.03	0.0064	1	0	100
k_2	0.02896	20.69	0.0014	1	0	100
k_3	0	0	0.0172	1	0	100
k_4	0	0	0.0026	1	0	100
k_5	0.00228	6.164	0.00037	1	0	100
k_6	0.1887	12.18	0.0155	1	0	100
k_e	0.000995	0.905	0.0011	1	0	100
D	5.33	0.148	35.92	1	0	100
t_{\max}	8.84	0.302	29.32	1	0	100
k	0.03	0.6	0.0444	1	0.6	3.4
y_b^{CB}	7.07	1.8	3.93	1	0.2	1.8

Tabella A.2: Stime parametri in fase di diagnosi ottenute per il paziente B.

(a) VWD Vicenza						
	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.09644	1.651	0.0584	1	0	100
k_1	0.0559	1.592	0.0351	1	0	100
k_2	0.0432	0.2805	0.154	1	0	100
k_3	9.017×10^{-8}	28.36	3.18×10^{-9}	1	0	100
k_4	0.0804	3.407	0.0236	1	0	100
k_5	0	0	0.00136	1	0	100
k_6	0.1719	1.938	0.0887	1	0	100
k_e	0.007688	0.9743	0.00789	1	0	100
D	29.45	0.4933	59.7	1	0	100
t_{\max}	26.72	0.8284	32.25	1	0	100
k	0.582	0.8094	0.719	1	0.6	1.5
y_b^{CB}	3.656	0.6553	5.58	1	0.2	1.8

(b) VWD 2B						
	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.2542	3.258	0.078	1	0	100
k_1	0.005788	0.3466	0.0167	1	0	100
k_2	0.01024	2.011	0.00509	1	0	100
k_3	0	0	0.0023	1	0	100
k_4	0	0	0.00439	1	0	100
k_5	0	0	5.54×10^{-5}	1	0	100
k_6	0.04242	0.2828	0.15	1	0	100
k_e	0.003336	1.192	0.0028	1	0	100
D	15.44	0.3433	44.98	1	0	100
t_{\max}	13.74	0.4522	30.39	1	0	100
k	0.1832	0.8	0.229	1	0.8	1.2
y_b^{CB}	14.94	1.486	10.05	1	0.2	1.8

(c) VWD 2A						
	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.2192	2.248	0.0975	1	0	100
k_1	0.004428	0.6919	0.0064	1	0	100
k_2	0.05532	39.52	0.0014	1	0	100
k_3	0	0	0.0172	1	0	100
k_4	0.1607	61.82	0.0026	1	0	100
k_5	4.241×10^{-5}	0.1146	0.00037	1	0	100
k_6	3.279×10^{-6}	0.0002115	0.0155	1	0	100
k_e	0.001207	1.097	0.0011	1	0	100
D	23.52	0.6548	35.92	1	0	100
t_{\max}	13.07	0.4458	29.32	1	0	100
k	0.02664	0.6	0.0444	1	0.6	3.4
y_b^{CB}	7.074	1.8	3.93	1	0.2	1.8

Tabella A.3: Stime parametri in fase di diagnosi ottenute per il paziente G.

(a) VWD Vicenza

	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.0593	1.015	0.0584	1	0	100
k_1	0.07842	2.234	0.0351	1	0	100
k_2	0.0775	0.5032	0.154	1	0	100
k_3	1.492×10^{-8}	4.691	3.18×10^{-9}	1	0	100
k_4	0	0	0.0236	1	0	100
k_5	0.001372	1.009	0.00136	1	0	100
k_6	0.3508	3.954	0.0887	1	0	100
k_e	0.007778	0.9858	0.00789	1	0	100
D	183.2	3.069	59.7	1	0	100
t_{\max}	12.41	0.3848	32.25	1	0	100
k	0.5911	0.8222	0.719	1	0.6	1.5
y_b^{CB}	2.515	0.4508	5.58	1	0.2	1.8

(b) VWD 2B

	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	3.319	42.55	0.078	1	0	100
k_1	0.005456	0.3267	0.0167	1	0	100
k_2	0.007876	1.547	0.00509	1	0	100
k_3	0	0	0.0023	1	0	100
k_4	0	0	0.00439	1	0	100
k_5	0.002763	49.87	5.54×10^{-5}	1	0	100
k_6	0.07843	0.5229	0.15	1	0	100
k_e	0.00314	1.122	0.0028	1	0	100
D	1.052	0.0234	44.98	1	0	100
t_{\max}	1.868	0.06147	30.39	1	0	100
k	0.1832	0.8	0.229	1	0.8	2.2
y_b^{CB}	11.46	1.141	10.05	1	0.2	1.8

(c) VWD 2A

	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.3035	3.113	0.0975	1	0	100
k_1	0.01306	2.041	0.0064	1	0	100
k_2	0.03703	26.45	0.0014	1	0	100
k_3	0	0	0.0172	1	0	100
k_4	0	0	0.0026	1	0	100
k_5	0.0007201	1.946	0.00037	1	0	100
k_6	0.0304	1.961	0.0155	1	0	100
k_e	0.000853	0.7754	0.0011	1	0	100
D	9.197	0.256	35.92	1	0	100
t_{\max}	12.9	0.4401	29.32	1	0	100
k	0.02664	0.6	0.0444	1	0.6	3.4
y_b^{CB}	7.074	1.8	3.93	1	0.2	1.8

Tabella A.4: Stime parametri in fase di diagnosi ottenute per il paziente M.

(a) VWD Vicenza						
	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.04575	0.7834	0.0584	1	0	100
k_1	0.141	4.016	0.0351	1	0	100
k_2	0.0853	0.5539	0.154	1	0	100
k_3	2.106×10^{-8}	6.623	3.18×10^{-9}	1	0	100
k_4	0.04352	1.844	0.0236	1	0	100
k_5	0.0007789	0.5727	0.00136	1	0	100
k_6	0.1491	1.681	0.0887	1	0	100
k_e	0.008382	1.062	0.00789	1	0	100
D	143.6	2.405	59.7	1	0	100
t_{\max}	22.69	0.7037	32.25	1	0	100
k	0.5758	0.8008	0.719	1	0.6	1.5
y_b^{CB}	2.723	0.4879	5.58	1	0.2	1.8

(b) VWD 2B						
	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.1343	1.722	0.078	1	0	100
k_1	0.006849	0.4101	0.0167	1	0	100
k_2	0.009304	1.828	0.00509	1	0	100
k_3	0	0	0.0023	1	0	100
k_4	0	0	0.00439	1	0	100
k_5	0.001964	35.46	5.54×10^{-5}	1	0	100
k_6	0.07172	0.4781	0.15	1	0	100
k_e	0.003175	1.134	0.0028	1	0	100
D	25.25	0.5614	44.98	1	0	100
t_{\max}	22.44	0.7383	30.39	1	0	100
k	0.1832	0.8	0.229	1	0.8	2.2
y_b^{CB}	11.95	1.189	10.05	1	0.2	1.8

(c) VWD 2A						
	Stime	Normalizzati	θ_0	Guess	Low bound	Up bound
k_0	0.05825	0.5975	0.0975	1	0	100
k_1	0.01268	1.98	0.0064	1	0	100
k_2	0.04554	32.53	0.0014	1	0	100
k_3	0	0	0.0172	1	0	100
k_4	0	0	0.0026	1	0	100
k_5	0.001398	3.778	0.00037	1	0	100
k_6	0.01946	1.255	0.0155	1	0	100
k_e	0.001021	0.9278	0.0011	1	0	100
D	21.21	0.5906	35.92	1	0	100
t_{\max}	55.22	1.883	29.32	1	0	100
k	0.02664	0.6	0.0444	1	0.6	3.4
y_b^{CB}	7.074	1.8	3.93	1	0.2	1.8

Appendice B

Stime parametriche ottenute con il set riparametrizzato

In questa Appendice vengono riportate le stime parametriche, i valori nominali, le stime preliminari di riferimento, le matrici di varianza-covarianza e di correlazione e tutte le statistiche in uscita dalle stime effettuate con la riparametrizzazione secondo il §4.6.

I diversi contributi di Φ^{ML} sono riportati separatamente in termini:

- costanti relativi al modello di varianza scelto;
- variabili in riferimento ai dati sperimentali a disposizione;
- *SWR*, residui sulla differenza tra i profili predetti dalla stima e i punti sperimentali.

Le tabelle che seguono fanno riferimento ad ogni pool di dati, rispettivamente:

- Tabella B.1 soggetti 0;
- Tabella B.2 soggetti non-0;
- Tabella B.3 pazienti VWD Vicenza;
- Tabella B.4 pazienti VWD 2B;
- Tabella B.5 pazienti VWD 2A;

Per ciascuna sottotabella i valori evidenziati in grassetto significano: (a) il superamento del test t-value al 95% di confidenza, (b) il superamento del test χ^2 , mentre nelle sottotabelle (d) si evidenzia la correlazione dei parametri.

Si ricorda che i fattori di normalizzazione adottati per ogni pool di dati θ_0 , sono le stime preliminari effettuate nel Capitolo 3.

Tabella B.1: Stime parametri soggetti 0 e relative statistiche.

(a) Stime

	Stime	Normalizzati	Nominali	θ_0	t-value	σ
k_0	0.0480	1.063	3000	0.0574	8.5	0.062
k_1	0.0108	1.31	0.00824	0.00824	6×10^{-8}	$1.1 \times 10^{+7}$
k_2	0.0537	1.192	0.0451	0.0451	2.9×10^{-7}	$2 \times 10^{+6}$
k_3	0.0041	1.02	0.00402	0.00402	0.32	1.6
k_4	0.001256	1.055	0.00119	0.00119	2.5×10^{-9}	$2 \times 10^{+8}$
k_5	7.175×10^{-5}	0.8527	8.414×10^{-5}	8.414×10^{-5}	0.21	2
k_6	0.0237	0.9967	0.0238	0.0238	4.9×10^{-8}	$1 \times 10^{+7}$
k_e	1.507×10^{-3}	1.076	0.14	0.00149	7	0.076
				Ref. t-value	1.6899	

(b) Residui

	Costanti	Variabili	SWR
y_{AG}	9.189	10.99	8.619
y_{CB}	9.189	10.99	9.072
TOT	18.379	21.972	17.69
Φ_{TOT}^{ML}			58.041
χ^2	23.059	Ref. χ^2 95% conf	49.802

(c) Matrice varianza-covarianza

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	0.00383							
k_1	-2.51×10^4	1.15×10^{14}						
k_2	4570	-2.09×10^{13}	4.2×10^{12}					
k_3	0.0204	-5.22×10^6	1.05×10^6	2.41				
k_4	-3.78×10^5	1.73×10^{15}	-3.47×10^{14}	-3.16×10^7	4.17×10^{16}			
k_5	0.0741	-5.2×10^6	1.04×10^6	-0.0482	-1.25×10^8	3.85		
k_6	1.9×10^4	-8.73×10^{13}	1.75×10^{13}	1.6×10^6	-2.1×10^{15}	6.77×10^6	1×10^{14}	
k_e	-0.00464	2.89×10^4	-5790	-0.0276	6.94×10^5	-0.0927	-3.32×10^4	0.0058

(d) Matrice correlazione

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	1							
k_1	-0.0378	1						
k_2	0.036	-0.952	1					
k_3	0.212	-0.313	0.329	1				
k_4	-0.0299	0.79	-0.83	-0.0998	1			
k_5	0.61	-0.247	0.259	-0.0158	-0.313	1		
k_6	0.0307	-0.812	0.853	0.103	-1.03	0.344	1	
k_e	-0.984	0.0354	-0.0371	-0.234	0.0447	-0.621	-0.0435	1

Tabella B.2: Stime parametri soggetti non-0 e relative statistiche.

(a) Stime

	Stime	Normalizzati	Nominali	θ_0	t-value	σ
k_0	0.05641	1.457	6000	0.0413	7.9	0.084
k_1	9.962×10^{-9}	1.48	6.73×10^{-9}	6.73×10^{-9}	2.2×10^{-13}	$3.1 \times 10^{+12}$
k_2	0.0535	0.64	0.0836	0.0836	1.2×10^{-6}	$2.4 \times 10^{+5}$
k_3	0.01846	0.6409	0.0288	0.0288	0.48	0.61
k_4	0.0004712	0.3228	0.00146	0.00146	8.1×10^{-16}	$1.8 \times 10^{+14}$
k_5	0.000143	0.9796	0.000146	0.000146	0.96	0.47
k_6	0.1297	1.804	0.0719	0.0719	2.1×10^{-13}	$4 \times 10^{+12}$
k_e	0.0006455	0.9765	0.0661	0.000661	7.2	0.062
				Ref. t-value	1.7827	

(b) Residui

	Costanti	Variabili	SWR
YAG	9.189	10.99	2.804
YCB	9.189	10.99	2.822
TOT	18.378	21.972	5.625
Φ_{TOT}^{ML}			45.976
χ^2	35.381	Ref. χ^2 95% conf	21.026

(c) Matrice varianza-covarianza

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	0.00711							
k_1	-2.3×10^9	9.47×10^{24}						
k_2	185	-7.62×10^{17}	5.62×10^{10}					
k_3	0.00313	6.29×10^{11}	-4.63×10^4	0.378				
k_4	-1.38×10^{11}	5.67×10^{26}	-4.18×10^{19}	4.6×10^{13}	3.33×10^{28}			
k_5	0.0237	-7.77×10^{11}	5.73×10^4	-0.0771	-4.56×10^{13}	0.219		
k_6	2.8×10^9	-1.15×10^{25}	8.49×10^{17}	-9.35×10^{11}	-6.75×10^{26}	1.12e+012	1.59×10^{25}	
k_e	-0.00521	1.5×10^9	-110	-0.0026	8.78×10^{10}	-0.0176	-2.07×10^9	0.00387

(d) Matrice correlazione

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	1							
k_1	-0.00886	1						
k_2	0.00926	-1	1					
k_3	0.0604	0.333	-0.318	1				
k_4	-0.00896	1	-0.968	0.411	1			
k_5	0.601	-0.54	0.517	-0.268	-0.534	1		
k_6	0.00831	-0.938	0.898	-0.381	-0.928	0.598	1	
k_e	-0.994	0.00783	-0.00749	-0.0681	0.00774	-0.605	-0.00835	1

Tabella B.3: Stime parametri pazienti affetti VWD Vicenza e relative statistiche.

(a) Stime

	Stime	Normalizzati	Nominali	θ_0	t-value	σ
k_0	0.07505	1.322	300	0.0584	7.3	0.083
k_1	0.01814	0.5169	0.0351	0.0351	9.3×10^{-8}	$2.6 \times 10^{+6}$
k_2	0.1329	0.8632	0.154	0.154	1.9×10^{-7}	$2.1 \times 10^{+6}$
k_3	1.197×10^{-9}	0.3764	3.18×10^{-9}	3.18×10^{-9}	2.5×10^{-7}	$7 \times 10^{+5}$
k_4	0.02694	1.142	0.0236	0.0236	1.7×10^{-7}	$3.2 \times 10^{+6}$
k_5	0.001389	1.021	0.00136	0.00136	0.49	0.96
k_6	0.1467	1.654	0.0887	0.0887	1×10^{-7}	$7.6 \times 10^{+6}$
k_e	0.007567	0.9609	0.315	0.00789	5.8	0.076
				Ref. t-value	1.7827	

(b) Residui

	Costanti	Variabili	SWR
YAG	9.189	10.99	0.3538
YCB	9.189	10.99	1.46
TOT	18.379	21.972	1.8135
Φ_{TOT}^{ML}			42.164
χ^2	3.627	Ref. χ^2 95% conf	21.026

(c) Matrice varianza-covarianza

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	0.00694							
k_1	546	6.57×10^{12}						
k_2	-466	-5.61×10^{12}	4.38×10^{12}					
k_3	8050	-1.9×10^{11}	1.48×10^{11}	4.93×10^{11}				
k_4	820	9.87×10^{12}	-7.71×10^{12}	-2.17×10^{11}	1×10^{13}			
k_5	0.0359	-1.57×10^6	1.23×10^6	-8.38×10^4	-1.59×10^6	0.926		
k_6	-1710	-2.05×10^{13}	1.6×10^{13}	4.5×10^{11}	-2.08×10^{13}	4.22×10^6	5.78×10^{13}	
k_e	-0.00618	-818	639	-8560	-829	-0.0333	2300	0.00573

(d) Matrice correlazione

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	1							
k_1	0.00256	1						
k_2	-0.00267	-1	1					
k_3	0.138	-0.105	0.101	1				
k_4	0.00311	1	-1	-0.0975	1			
k_5	0.448	-0.637	0.609	-0.124	-0.523	1		
k_6	-0.00269	-1	1	0.0843	-0.865	0.577	1	
k_e	-0.98	-0.00422	0.00403	-0.161	-0.00346	-0.457	0.004	1

Tabella B.4: Stime parametri pazienti affetti da VWD 2B e relative statistiche.

(a) Stime

	Stime	Normalizzati	Nominali	θ_0	t-value	σ
k_0	0.08322	1.162	1000	0.078	5.9	0.09
k_1	0.01665	0.9972	0.0167	0.0167	1×10^{-7}	$4.4 \times 10^{+6}$
k_2	0.005136	1.009	0.00509	0.00509	6.5×10^{-8}	$7.1 \times 10^{+6}$
k_3	0.001402	0.6096	0.0023	0.0023	0.018	16
k_4	0.002426	0.5526	0.00439	0.00439	2.5×10^{-9}	$1 \times 10^{+8}$
k_5	7.438×10^{-5}	1.343	5.54×10^{-5}	5.54×10^{-5}	0.03	20
k_6	0.1613	1.075	0.15	0.15	1.6×10^{-7}	$3.1 \times 10^{+6}$
k_e	0.002865	1.023	0.112	0.0028	4.8	0.098
				Ref. t-value	1.7827	

(b) Residui

	Costanti	Variabili	SWR
Y_{AG}	9.189	10.99	8.625
Y_{CB}	9.189	10.99	0.3817
TOT	18.379	21.972	9.0068
Φ_{TOT}^{ML}			49.358
χ^2	18.014	Ref. χ^2 95% conf	21.026

(c) Matrice varianza-covarianza

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	0.00806							
k_1	-745	1.95×10^{13}						
k_2	2450	-6.41×10^{13}	5.04×10^{13}					
k_3	-0.0116	-4.86×10^7	3.82×10^7	251				
k_4	-2.75×10^4	7.2×10^{14}	-5.66×10^{14}	-9.83×10^8	1.05×10^{16}			
k_5	0.373	1.2×10^7	-9.44×10^6	-101	1.76×10^8	417		
k_6	804	-2.11×10^{13}	1.66×10^{13}	2.88×10^7	-3.09×10^{14}	-9.43×10^6	9.86×10^{12}	
k_e	-0.00871	3450	-2710	0.0108	5.06×10^4	-0.403	-1620	0.00951

(d) Matrice correlazione

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	1							
k_1	-0.00188	1						
k_2	0.00384	-1	1					
k_3	-0.00819	-0.695	0.34	1				
k_4	-0.00298	1	-0.776	-0.605	1			
k_5	0.203	0.133	-0.0651	-0.313	0.0839	1		
k_6	0.00285	-1	0.743	0.579	-0.957	-0.147	1	
k_e	-0.995	0.00801	-0.00392	0.00702	0.00505	-0.202	-0.00528	1

Tabella B.5: Stime parametri pazienti affetti da VWD 2A e relative statistiche.

(a) Stime

	Stime	Normalizzati	Nominali	θ_0	t-value	σ
k_0	0.09161	1.118	3000	0.0975	8.6	0.06
k_1	0.0068	1.062	0.0064	0.0064	5.4×10^{-8}	$9.1 \times 10^{+6}$
k_2	0.00121	0.8642	0.0014	0.0014	1.1×10^{-8}	$3.5 \times 10^{+7}$
k_3	0.01776	1.032	0.0172	0.0172	2.4	0.2
k_4	0.003254	1.252	0.0026	0.0026	1.3×10^{-8}	$4.4 \times 10^{+7}$
k_5	0.000391	1.057	0.00037	0.00037	0.13	3.8
k_6	0.015	0.9678	0.0155	0.0155	5.6×10^{-8}	$8 \times 10^{+6}$
k_e	0.001093	0.9933	0.044	0.0011	7.8	0.058
				Ref. t-value	1.7827	

(b) Residui

	Costanti	Variabili	SWR
y_{AG}	9.189	10.99	77.37
y_{CB}	9.189	10.99	0.6577
TOT	18.379	21.972	78.029
Φ_{TOT}^{ML}			118.38
χ^2	156.06	Ref. χ^2 95% conf	21.026

(c) Matrice varianza-covarianza

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	0.00355							
k_1	-1370	8.23×10^{13}						
k_2	6260	-3.76×10^{14}	1.21×10^{15}					
k_3	0.00217	-1.42×10^6	4.56×10^6	0.0393				
k_4	-7240	4.35×10^{14}	-1.4×10^{15}	-7.4×10^6	1.94×10^{15}			
k_5	0.00993	-8.2×10^6	2.63×10^7	0.0308	-3.66×10^7	14.2		
k_6	1220	-7.3×10^{13}	2.34×10^{14}	1.24×10^6	-3.26×10^{14}	5.21×10^6	6.4×10^{13}	
k_e	-0.00346	2400	-7710	-0.00228	1.07×10^4	-0.00976	-2110	0.0034

(d) Matrice correlazione

	k_0	k_1	k_2	k_3	k_4	k_5	k_6	k_e
k_0	1							
k_1	-0.00254	1						
k_2	0.00303	-1	1					
k_3	0.184	-0.791	0.663	1				
k_4	-0.00276	1	-0.912	-0.847	1			
k_5	0.0443	-0.24	0.201	0.0413	-0.22	1		
k_6	0.00255	-1	0.843	0.783	-0.924	0.173	1	
k_e	-0.997	0.00454	-0.00381	-0.197	0.00417	-0.0445	-0.00452	1

Riferimenti bibliografici

- Asprey, S. P. e S. Macchietto (2000). Statistical tools for optimal dynamic model building. *Comput. Chem. Eng.*, **24**, 1261-1267.
- Bard, Y. (1977). *Nonlinear parametric estimation*. Academic Press, New York (U.S.A.).
- Box, G. E. P. e H. L. Lucas (1959). Design of experiments in non-linear situations. *Biometrika*, **46**, 77-90.
- Buzzi-Ferraris, G. e P. Forzatti (1983). A New Sequential Experimental Design Procedure for Discriminating Among Rival Models. *Chem. Eng. Sci.*, **38**, 225.
- Carson, E. e C. Cobelli (2001). *Modelling methodology for physiology and medicine*. Academic Press, San Diego (U.S.A.).
- Casonato A., E. Pontara, A. Bertomoro, S. Zucchetto, P. Zerbinati e A. Girolami (1996). Abnormal collagen binding activity of 2A von Willebrand factor: Evidence that the defect depends only on the lack of large multimers. *Jour. Lab.Clin. Med*, **129-2**, 251-259.
- Casonato A., L. Gallinaro, M. G. Cattini, E. Pontara, R. Padrini, A. Bertomoro, V. Daidone e A. Pagnan (2010). Reduced survival of type 2B von Willebrand factor, irrespective of large multimer representation or thrombocytopenia. *Haematologica*, **95(8)**, 1366-1372.
- Cobelli, C. (2007). Identificazione di modelli e identificabilità a priori. *Appunti per il Corso di Modelli e Controllo di Sistemi Biologici 2*, DEI, Università di Padova
- Chen, B. H. e S. P. Asprey (2003). On the design of optimally informative dynamic experiments for model discrimination in multi response non-linear situations. *Ind. Eng. Chem. Res.*, **42**, 1379-1390.
- De Luca, R. (2012). *Tecniche di riparametrizzazione nella progettazione ottimale di esperimenti per l'identificazione di modelli*. Tesi di Laurea Magistrale in Ingegneria Chimica e dei Processi Industriali, DIPIC, Università di Padova.

- Emery, A. F. (2001). Using the concept of information to optimally design experiments with uncertain parameters. *ASME J. Heat Transfer*, **123**, 593-600.
- Fisher, R. A. (1935). *The design of the experiments*. Oliver & Boyd, Edinburgh (U.K.)
- Galvanin, F., M. Barolo e F. Bezzo (2009). Online model-based redesign of experiments for parametr estimation in dynamic systems. *Ind. Eng. Chem. Res.*, **48**, 4415-4427.
- Galvanin, F. (2010). *Optimal model-based design of experiments in dynamic systems: novel techniques and unconventional applications*. Tesi di Dottorato in Ingegneria Chimica, DIPIC, Università di Padova.
- Galvanin, F. (2012a). Report 1 DII, Università di Padova.
- Galvanin, F. (2012b). Report 2 DII, Università di Padova.
- Galvanin, F. (2012c). Report 2bis DII, Università di Padova.
- Galvanin, F. (2012d). Report 3 DII, Università di Padova.
- Goodeve, A. C. (2010). The genetic basis of von Willebrand disease. *Blood Reviews*, **24**, 123-134.
- Hoffman, M. e D. M. Monroe III (2001). A Cell-based Model of Hemostasis. *Thromb. Haemost.*, **85**, 958-965.
- Lillicrap, D. (2007). Von Willebrand disease-Phenotype versus genotype: Deficiency versus disease. *Thrombosis Research*, **120**, S11-S16.
- Menache, D., D.L. Aronson, F. Darr, R.R. Montgomery, J.C. Gill, C.M. Kessler, J.M. Lusher, P.D. Phatak, A.D. Shapiro, A.R. Thompson e G.C White II (1996). Pharmacokinetics of von Willebrand factor VIIIc in patients with severe von Willebrand disease (type 3 VWD): estimation of the rate of factor VIIIc synthesis. Cooperative Study Groups. *Br. J. Haematol.*, **94**(4), 740-745.
- N.H.L.B.I. (2007). *The Diagnosis, Evaluation and Management of von Willebrand Disease*. NIH Publication No. 8-5832, U.S. Department of Health and Human Services (U.S.A.), available at: <http://www.nhlbi.nih.gov/guidelines/vwd>

- PSE gPROMS[®] (2010). *Model Validation Guide* (Release v3.2.1). Process Systems Enterprise Limited, London (U.K.).
- Rodeghiero, F., G. Castaman e E. Dini (1987). Epidemiological investigation of the prevalence of von Willebrand's disease. *Blood*, **69**, 454-459.
- Sadler, J. E. (1998). Biochemistry and genetics of von Willebrand factor. *Annu. Rev. Biochem.*, **67**, 395-424.
- Sadler, J. E. (2005). von Willebrand factor: two sides of a coin. *Jour. Throm. Haemos.*, **3**, 1702-1709.
- Shaw, B. M. (1999). *Statistical issue in kinetic modeling of gas-phase ethylene copolymerisation*. Ph.D Thesis, Queen's University, Canada.
- Söderström, T. e P. Stoica (1989). *System identification*. Prentice Hall, New York (U.S.A.).
- Werner, E. J., E. H. Bronxson, E. L. Tucker, D. S. Giroux, J. Shults e T. C. Abshire (1993). Prevalence of von Willebrand disease in children: a multiethnic study. *Jour. Pediatr.*, **123**, 893-898.
- Ye, X., C. Y. Suen, M. Cheriet e E. Wang (1999). A Recent Development in Image Analysis of Electrophoresis Gels. *Vision Interface*, 19-21 May (Canada), 432-438.
- Zullo, L. (1991). *Computer aided design of experiments. An engineering approach*. PhD. Thesis, University of London, U.K..

Ringraziamenti

Un ringraziamento alla **Dott.ssa Sandra Casonato** del Dipartimento di Scienze Cardiologiche, Toraciche e Vascolari dell'Università degli Studi di Padova e al **Prof. Roberto Padrini** del Dipartimento di Medicina Clinica e Sperimentale dell'Università degli studi di Padova per i dati forniti e per l'apporto tecnico e professionale prestato.

Grazie al **Prof. Fabrizio Bezzo** al **Dr. Federico Galvanin** e al **Prof. Massimiliano Barolo** per la costante supervisione ed i preziosissimi consigli durante questi mesi di lavoro.