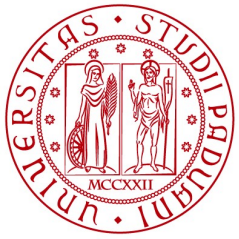


Università degli Studi di Padova

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Ingegneria Industriale

Scuola di dottorato di ricerca in INGEGNERIA INDUSTRIALE
Indirizzo: INGEGNERIA CHIMICA, DEI MATERIALI E DELLA PRODUZIONE
Ciclo XXVII

A Model-Based investigation on the effect of Light on Microalgae Growth: Focusing on Photoproduction, Photoregulation and Photoinhibition

Direttore della Scuola : Ch.mo Prof. Paolo Colombo
Coordinatore di indirizzo: Ch.mo Prof. Enrico Savio
Supervisore : Ch.mo Prof. Fabrizio Bezzo

Dottorando: Andrea Bernardi

Foreword

The fulfilment of the research project presented in this Thesis has involved the financial and intellectual support of many people, to whom the author is most grateful. Most of the research activity that led to the results and achievements summarised in the Thesis has been developed at the Department of Industrial Engineering of the University of Padova (DII), under the supervision of Prof. Fabrizio Bezzo. Part of the work has been carried out at the Centre for Process Systems Engineering, Chemical Engineering Department, Imperial College London (UK), under the supervision of Dr. Benoit Chachuat.

The realisation of this study has been made possible thanks to the financial support of Fondazione Cariparo under Progetto Dottorati di Ricerca 2012.

All the material presented in this Thesis is original, unless explicit references provided by the author. The full list of publications drawn from this research project is reported below.

Articles published in international journals

Nikolaou, A., Bernardi, A., Meneghesso, M., Bezzo, F., Morosinotto, T., B. Chachuat (2015). A model of chlorophyll fluorescence in microalgae integrating photoproduction, photoinhibition and photoregulation. *Journal of biotechnology*, **194**, 91-99.

Bernardi, A., G. Perin,F. Galvanin, T. Morosinotto, F. Bezzo (2014). An identifiable state model to describe light intensity influence on microalgae growth., *Industrial & Engineering Chemistry Research*, **53**, 6738-6749.

Bernardi, A., Giarola, S., F. Bezzo (2013). Spatially explicit multi-objective optimisation for the strategic design of first and second generation biorefineries including carbon and water footprints. *Industrial & Engineering Chemistry Research*, **52**, 7170-7180.

Bernardi, A., Giarola, S., F. Bezzo (2012). Optimizing the economics and the carbon and water footprints of bioethanol supply chains. *Biofuels, Bioprodcts and Biorefineries*, **6**, 656-672.

Publications/Abstracts in conference proceedings

Nikolaou, A., Bernardi, A., Bezzo, F., Morosinotto, T., B. Chachuat (2014). Dynamic modeling of photoproduction, photoregulation and photoinhibition in microalgae using

chlorophyll fluorescence. Presented at: *AIChE 2014 Annual Meeting*, November 16-21, 2014, Atlanta, USA.

Nikolaou, A., Bernardi, A., Bezzo, F., Morosinotto, T., B. Chachuat (2014). A dynamic model of photoproduction, photoregulation and photoinhibition in microalgae using chlorophyll fluorescence. Presented at: *19th IFAC World Congress*, Cape Town (South Africa), 24-29 August 2014.

Bernardi, A., F. Galvanin, G. Perin, T. Morosinotto, F. Bezzo (2014) Modelling the effect of light intensity in microalgae growth. Presented at: *AIChE Annual Meeting*, San Francisco (U.S.A.), 3-8 November.

Ortiz Gutierrez, R., Penazzi, S., Bernardi, A., Giarola, S., F. Bezzo (2013). A spatially-explicit approach to the design of ethanol supply chains considering multiple technologies and carbon trading effects. In: *Computer-Aided Chemical Engineering 32, 23rd European Symposium on Computer Aided Process Engineering* (A. Kraslawski and I. Turunen, Eds.), Elsevier, Amsterdam (The Netherlands) 643-648.

Bernardi, A., Giarola, S., F. Bezzo (2012). A framework for water footprint optimization along the bioethanol supply chain. *Computer-Aided Chemical Engineering 31, 11th International Symp. on Process Systems Engineering* (I.A. Karimi, R. Srinivasan Eds.), Elsevier, Amsterdam (The Netherlands), 1372-1376.

Oral presentations

Bernardi, A., Giarola, S., Bezzo, F., 2012. Pianificazione strategica di una filiera di produzione di etanolo: minimizzazione di emissioni e dell?impatto idrico. In: *Convegno GRICU 2012*, Montesilvano (ITALY), 16-19 September.

Articles submitted for pubblication in international journals or conference proceedings

Meneghesso, A., Bernardi, A., Nikolaou, A., Perin, G., Chachuat, B., Bezzo, F., T. Morosinotto. Computational models to improve light use efficiency in microalgae culture. Submitted to: *Journal of Biotechnology*

Bernardi, A., Nikolaou, A., Meneghesso, A., Bezzo, F., Morosinotto, T., B. Chachuat. Using fluorescence measurements to model key phenomena in microalgae photosynthetic mechanisms. Submitted to: *Chemical Engineering Transactions*.

Bernardi, A., Nikolaou, A., Meneghesso, M., Bezzo, F., Morosinotto, T., B. Chachuat; A framework for dynamic modelling of photosynthesis-irradiance curves in microalgae. Submitted to: *12th PSE and 25th ESCAPE Joint Event*, Copenhagen, Denmark. 31 May - 4 June 2015.

Abstract

Biofuels derived from microalgae may represent a key source for alternative energy vectors. Moreover, microalgae exhibit a great potential for sustainable production of a wide range of commodities and value-added products, including cosmetics, pharmaceuticals and nutraceuticals, which makes them suitable for biorefinery applications. Their high productivity and their ability to accumulate large amounts of lipids, along with their independence from arable land, put them in a competitive position with respect to traditional oil crops. However, the economical and energetic sustainability of large scale microalgae cultivation for biodiesel production are still debated. The most optimistic previsions are in fact based on gross estimates of productivity, derived by extrapolation of laboratory-scale data.

Therefore, the development of reliable mathematical models that are capable of quantitative predictions of the behaviour of large-scale outdoor microalgae culture is of paramount importance. Such models prove especially useful in identifying which parameters have the largest impact on productivity, thereby providing a means for enhancing the growth conditions through design and operational changes. Moreover, accurate forecasts of microalgal growth in the outdoor conditions can lead to a better understanding of the real potential of microalgae-based biofuels.

This Thesis aims of investigating the complex effect of light in the photosynthetic apparatus activity, and its effect on microalgae growth. The work presented in this Thesis follows two general lines. The first contribution has been to propose a general approach for model development. The proposed methodology guides the modelling effort in order to assure both an accurate representation of the calibration data, but most importantly also the identifiability of the model. The identifiability of a model, i.e. the possibility to estimate in accurate and reliable way its parametric set, is in fact, a necessary property for the model to be confidently used in process scale-up and optimization. The proposed methodology has been successfully applied to growth and fluorescence data of the sea water alga Nannochloropsis Salina.

A second contribution is concerned with Pulsed Amplitude Modulation (PAM) fluorometry. A dynamic model of chlorophyll fluorescence has been developed. The model integrates photoproduction, photoregulation and photoinhibition processes in a semi-mechanistic way. The model has been calibrated against fluorescence data of a sample of the microalga Nannochloropsis gaditana. The proposed fluorescence model is capable of quantitative prediction of the state of the photosynthetic apparatus of microalgae in terms of their open, closed and damaged reaction centres under variable light conditions. Two promising application of the fluorescence model have also been analysed: (i) the model has been used for the prediction

of photosynthesis rate versus irradiance (PI)-response curves based on PAM fluorometry; and (ii) a model based experiment design (MBDoE) approach has been followed to define new information rich PAM protocols to further validate and refine the model structure.

Riassunto

Nonostante la crisi finanziaria, il consumo energetico negli ultimi anni è continuato ad aumentare, spinto principalmente dall'aumento della popolazione mondiale e dalla crescita economica dei paesi in via di sviluppo (in particolare Cina ed India). Oltre a preoccupazioni riguardanti la limitatezza delle riserve di energia fossile, il problema del riscaldamento globale rende sempre più evidente l'insostenibilità ambientale dell'attuale sistema di approvvigionamento di energia. Particolarmente critico è il settore dei trasporti, che attualmente è quasi del tutto dipendente dal petrolio. L'elevata densità energetica fornita dai combustibili liquidi, infatti, lascia poco spazio ad alternative quali i motori elettrici o a biogas, almeno per quanto riguarda il breve e medio periodo.

Uno dei principali vantaggi dei biocarburanti è precisamente quello di poter essere impiegati nei motori attualmente sul mercato in sostituzione ai combustibili fossili o miscelati ad essi e di poter sfruttare il sistema di distribuzione impiegato per i carburanti tradizionali. Tra le alternative disponibili per la produzione di biomassa idonea alla conversione in biocombustibili, le microalghe sono considerate una delle fonti di biomassa più promettenti. L'elevata produttività, l'indipendenza da terreni agricoli e la possibilità di utilizzare acque non potabili e non adatte alle coltivazioni (acqua di mare o salmastra), assieme alla capacità di alcune specie di microalghe di accumulare elevate quantità di olio, le rendono il candidato ideale. Tuttavia, la produzione su larga scala di microalghe in modo economicamente ed ambientalmente sostenibile richiede ancora un notevole lavoro di ricerca.

Le previsioni più ottimistiche, infatti, sono basate su estrapolazioni da dati ottenuti in laboratorio, dove le condizioni di crescita sono notevolmente diverse da quelle in un sistema di produzione su larga scala. La disponibilità di modelli matematici in grado di prevedere il comportamento delle colture microalgali in sistemi di coltivazione industriali è quindi di primaria importanza per progettare, simulare e ottimizzare i processi di produzione. Per essere sufficientemente robusto e affidabile quando utilizzato in predizione, un modello matematico deve basarsi per quanto possibile su considerazioni di tipo meccanicistico e in secondo luogo deve essere identificabile (ovvero deve essere possibile stimare in modo univoco il set di parametri che lo caratterizzano).

L'analisi di letteratura svolta sui modelli esistenti ha messo in luce alcune limitazioni che questa Tesi si propone di superare. Il primo limite individuato è che spesso i modelli letteratura sono, o eccessivamente semplificati, e quindi incapaci di descrivere in modo sufficientemente accurato il sistema considerato, o al contrario, estremamente complessi e quindi difficilmente identificabili. Di conseguenza, il primo obiettivo di questa Tesi (affrontato nel Capitolo 3) è

stato quello di proporre uno schema generale in grado di guidare lo sviluppo di nuovi modelli, assicurando due caratteristiche fondamentali: la prima è che i modelli devono essere in grado di rappresentare i dati in modo preciso, la seconda è che il modello sviluppato deve essere identificabile sulla base dei dati disponibili.

Un secondo limite evidenziato dall'analisi di letteratura riguarda invece l'utilizzo di misure di fluorescenza. Strumenti quali la fluorometria PAM (*Pulsed Amplitude Modulation*) rappresentano infatti lo stato dell'arte degli strumenti di analisi dei processi fotosintetici e hanno portato ad importanti scoperte in abito biologico negli ultimi 40 anni. Tuttavia, il loro utilizzo è spesso limitato ad analisi di tipo qualitativo e poco è stato fatto per lo sviluppo di modelli matematici in grado di fornire informazioni quantitative. Questo ha dato l'impulso alla seconda parte della Tesi (Capitoli 4-6) nella quale è stato sviluppato un modello dinamico in grado di rappresentare la fluorescenza della clorofilla in funzione dell'intensità luminosa. Il modello, tenendo conto dei processi di fotoproduzione, fotoregolazione e fotoinibizione, è in grado di fornire previsioni accurate del flusso di fluorescenza misurato in varie condizioni sperimentali e ha permesso la derivazione di relazioni quantitative tra le misure di fluorescenza e lo stato del fotosistema in termini di centri di reazione aperti, chiusi e inibiti. Inoltre, sono state studiate altre due importanti applicazioni del modello di fluorescenza. La prima riguarda la possibilità di predire la produttività fotosintetica basandosi su misure di fluorescenza, molto più veloci ed affidabili rispetto alle misure tradizionali. La seconda applicazione del modello riguarda invece l'utilizzo di tecniche di progettazione ottimale di esperimenti basate su modello (*Model based experiment design*, MBDoE) per la determinazione di esperimenti ottimali in termini di informazione contenuta in essi.

Il lavoro di Tesi è organizzato secondo il seguente schema concettuale.

Nel Capitolo 1 dopo aver inquadrato il panorama bibliografico di riferimento, vengono illustrate le principali limitazione degli attuali approcci modellistici usati per descrivere la crescita delle microalghe.

Nel Capitolo 2 vengono illustrati i principali strumenti matematici e sperimentali utilizzati nel proseguo della Tesi. Dopo aver presentato uno schema generale per lo sviluppo di un modello matematico che sia sufficientemente affidabile e robusto vengono discussi i concetti di identificabilità *a priori* e *a posteriori*; i fondamenti teorici delle tecniche MBDoE e il metodo di stima parametrica basato sulla massima verosimiglianza. Nella seconda parte del capitolo vengono introdotte le misure di fluorescenza della clorofilla fornendo alcune informazioni biologiche di base necessarie a comprendere le fondamenta teoriche di questo tipo di misure.

Nel Capitolo 3 viene sviluppato un modello in grado di rappresentare dei dati di crescita e di fluorescenza riguardanti la specie Nannochloropsis salina. Attraverso opportune tecniche di discriminazione è stato selezionato il modello più promettente tra dei candidati reperiti nella letteratura scientifica. Tale modello è stato ulteriormente modificato, da un lato per meglio rappresentare alcuni importanti fenomeni biologici, dall'altro per assicurare l'identificabilità

strutturale del modello stesso.

Il Capitolo 4 presenta un modello semi-meccanicistico della fluorescenza della clorofilla. Il modello considera i principali processi biologici che agiscono in tempi caratteristici compresi tra i millisecondi (fotoproduzione) e le ore (fotoinibizione). In particolare, vengono usate misure di fluorescenza PAM per sviluppare e calibrare il modello. Tra i vantaggi offerti da un modello matematico come quello proposto ci sono: (i) la possibilità di verificare le condizioni sotto le quali sono valide alcune ipotesi biologiche, correntemente utilizzate in letteratura; e (ii) stimare l'andamento del numero di centri di reazione inibiti durante l'esperimento di fluorescenza.

Il Capitolo 5 propone un'estensione del modello sviluppato nel Capitolo 4. In particolare viene introdotto nel modello il processo di fotoacclimatazione e un'espressione per il calcolo della produttività fotosintetica in termini di evoluzione di ossigeno (curve PI). La simulazione dinamica delle curve PI ha messo in luce l'importanza del protocollo usato durante gli esperimenti per misurare l'evoluzione di ossigeno. Questo è un risultato particolarmente significativo dal momento che spesso in letteratura viene trascurato l'effetto della dinamica e l'utilizzo di curve PI viene fatto ipotizzando che esse rappresentino uno stato stazionario. L'analisi dei risultati ottenuti mostra invece come l'assunzione di tale ipotesi possa causare una sovrastima della produttività e debba essere pertanto evitata.

Nel Capitolo 6 attraverso l'utilizzo di tecniche MBDoE la struttura del modello di fluorescenza viene ulteriormente raffinata e validata. Viene inoltre migliorata l'accuratezza della stima parametrica. Infine, un nuovo tipo di misure di fluorescenza, ottenute con un diverso tipo di fluorometro vengono utilizzate per caratterizzare meglio la dinamica della fotoproduzione.

Il capitolo 7 conclude la Tesi riassumendo i principali risultati raggiunti e indicando possibili sviluppi futuri per proseguire la ricerca sull'argomento.

Table of Contents

Foreword i

Abstract iii

Riassunto v

Table of Contents ix

Acronyms xiii

List of Symbols xv

List of Figures xviii

List of Tables xxii

1 Introduction 1
- 1.1 Energy outlook . . . 2
- 1.2 Climate change . . . 3
- 1.3 Renewables in transport sector . . . 6
- 1.4 Photosynthesis fundamentals . . . 8
- 1.5 Third generation biofuels from microalgae . . . 10
  - 1.5.1 Environmental challenges and opportunities of microalgae . . . 11
  - 1.5.2 Microalgae culture systems . . . 12
- 1.6 Models for microalgae growth . . . 14
- 1.7 Motivation of the work . . . 18
- 1.8 Thesis roadmap . . . 19

2 Materials and methods 21
- 2.1 Mathematical background . . . 21
  - 2.1.1 Developing a first-principles model . . . 21
  - 2.1.2 Identifiability analysis . . . 24
  - 2.1.3 Model based design of experiments . . . 25
  - 2.1.4 Parameter estimation . . . 27
- 2.2 Experimental background . . . 28
  - 2.2.1 Principle of chlorophyll fluorescence . . . 28
  - 2.2.2 Pulsed amplitude modulation protocols . . . 30
  - 2.2.3 Inference of fluorescence protocols: fluorescence indexes . . . 31
  - 2.2.4 Utilisation of PAM experiments . . . 32

3 An identifiable state model to describe light intensity influence on microalgae growth 35
3.1 Motivation . . . 35
3.2 Model developing approach . . . 37
3.3 Modelling approaches . . . 39
3.3.1 Eilers Peeters model . . . 40
3.3.2 Camacho Rubio model . . . 41
3.4 Experimental set-up and available data . . . 42
3.5 Model discrimination and preliminary parameter estimation . . . 44
3.5.1 Enhancing CRM . . . 46
3.6 Identifiability analysis . . . 48
3.6.1 Global identifiability analysis . . . 48
3.6.2 Sensitivity and correlation analysis . . . 50
3.6.3 Reparameterisation of Camacho Rubio model . . . 52
3.7 Results and discussion . . . 53
3.7.1 Model calibration . . . 53
3.7.2 Model validation . . . 55
3.8 Conclusion . . . 56

4 A model of chlorophyll fluorescence in microalgae integrating photoproduction, photoinhibition and photoregulation 59
4.1 Motivation . . . 60
4.2 Material and methods . . . 61
4.3 A dynamic model of fluorescence in microalgae . . . 62
4.3.1 Han model . . . 63
4.3.2 Accounting for photoregulation . . . 65
4.3.3 Properties of fluorescence model . . . 67
4.4 Results and discussion . . . 69
4.4.1 Model calibration . . . 69
4.4.2 Model analysis and validation . . . 71
4.5 Conclusions . . . 73

5 A framework for the dynamic modelling of PI curves in microalgae 75
5.1 Motivation . . . 75
5.2 Dynamic modelling of Photosynthesis-Irradiance curves . . . 76
5.2.1 Photoacclimation extension . . . 76
5.2.2 Photosynthesis rate . . . 78
5.3 Results and discussion . . . 78
5.3.1 Dynamic simulation of PI curves . . . 79
5.4 Coupling PI curves with fluorescence data for model calibration . . . 82
5.4.1 Case study: in silico PI curves experiment . . . 82
5.4.2 Case study: calibration with literature data . . . 84
5.5 Conclusion . . . 87

6 MBDoE approach for information rich PAM experiments 89
6.1 Motivation . . . 89
6.2 Antenna size measurements . . . 90
6.3 Model based design of experiment to improve the parameter estimation of fluorescence model . . . 92

6.3.1 Optimal experimental design 92
6.3.2 Practical model identifiability 94
6.3.3 Estimating recovery of inhibited PSUs 95
6.4 Model enhancing: two steps NPQ 95
6.5 Design of experiment for the complex NPQ model 99
6.5.1 Measurement error evaluation 100
6.5.2 Practical model identifiability 101
6.5.3 Estimating recovery of inhibited PSUs 102
6.6 Complex NPQ model validation 102
6.7 Conclusion 104

7 Conclusions and future perspectives **107**
7.1 Summary of thesis achievements 107
7.2 Future perspective 110

A Fluorescence quantum yield **113**
A.1 Alternative representation of the Light-Harvesting Complex (LHC)-reaction centre (RC) complex 113

Bibliography **114**

Acronyms

3PG 3-phosphoglycerate.

ASII functional antenna size.

CRM Camacho Rubio model.

Cyt b_6f cytochrome b_6f.

DCMU 3-(3,4-dichlorophenyl)-1,1-dimethylurea.

EPM Eilers and Peeters model.

FIM Fisher information matrix.

FRR Fast Repetition Rate.

G3P glyceraldehyde-3-phosphate.

GHG green house gases.

IEA International Energy Agency.

LCA life cycle assessment.

LHC Light-Harvesting Complex.

LHCI Light-Harvesting Complex of photosystem I.

LHCII Light-Harvesting Complex of photosystem II.

MBDoE model based design of experiments.

mCRM modified Camacho Rubio model.

MIMO multiple-input-multiple-output.

MTOE million tonnes of oil equivalent.

NER net energy ratio.

NPQ Non Photochemical Quenching.

PAM Pulse Amplitude Modulation.

PBR photobioreactors.

PC plastocyanin.

PI photosynthesis-irradiance.

PQ plastoquinone.

PSI Photosystem I.

PSII Photosystem II.

PSU Photo-Synthetic Unit.

RC reaction centre.

RCII reaction centre of photosystem II.

ROS reactive oxygen species.

RUBISCO ribulose-1,5-biphosphate carboxylase-oxygenase.

RuBP ribulose-1,5-biphosphate.

SGI structurally globally identifiable.

SLI structurally locally identifiable.

SNI structurally non-identifiable.

List of Symbols

σ total cross section [$\mathrm{m^2 g_{chl}^{-1}}$].

α_2 parameter that regulate acclimation in modified Camacho Rubio model [-].

η_D rate of basal termal decay relative to the rate of fluorescence [-].

ξ_F rate of NPQ adaptation [$\mathrm{s^{-1}}$].

Φ_f quantum yield of fluorescence [$\mu\mathrm{E}\,\mu\mathrm{E}^{-1}$].

Φ_f^A fluorescence quantum yield of a reaction centre in state A [-].

Φ_f^B fluorescence quantum yield of a reaction centre in state B [-].

Φ_f^C fluorescence quantum yield of a reaction centre in state C [-].

η_I rate of inhibition related quenching relative to the rate of fluorescence [-].

Φ_L scaled realised quantum yield of photosynthesis [-].

η_P rate of photoproduction relative to the rate of fluorescence [-].

Φ_p^A quantum yield of photosynthesis of an open reaction centre of the photosystem II [-].

λ_PAM gain parameter aligning the voltage output of a PAM fluorometer with the actual fluorescence flux [$\mathrm{V\,s}\,\mu\mathrm{E}^{-1}$].

Φ_PS2 realised quantum yield of photosynthesis [-].

σ_PSU optical cross section of a photosynthetic unit [$\mathrm{m^2}\mu\mathrm{E}^{-1}$].

η_{qE} rate of energy dependent quenching relative to the rate of fluorescence [-].

$\overline{\eta}_\mathrm{qE}^F$ maximum rate of fast energy dependent quenching relative to the rate of fluorescence [-].

$\overline{\eta}_\mathrm{qE}^{int}$ maximum rate of interaction energy dependent quenching relative to the rate of fluorescence [-].

$\overline{\eta}_\mathrm{qE}^S$ maximum rate of slow energy dependent quenching relative to the rate of fluorescence [-].

$\overline{\sigma}_{r,s}$ rs-th component of the inverse of variance-covariance matrix of experimental measurements.

α_S activity level of slow energy dependent quenching [-].

α_{SS} reference function for energy dependent quenching activity [-].

$\mathbf{\Sigma}_y$ variance-covariance matrix of experimental measurements.

A fraction of photosynthetic units in resting state according to Han model.

a_1 amount of photosynthetic units in resting state.

a_2 amount of photosynthetic units in activated state.

a_3 amount of photosynthetic units in inhibited state.

a_t total amount of photosynthetic units.

B fraction of photosynthetic units in activated state according to Han model.

C fraction of photosynthetic units in inhibited state according to Han model.

chl chlorophyll content in a sample [$\mathrm{g_{chl}\,m^{-3}}$].

F fluorescence flux usually measured in volts.

F' light-adapted realised fluorescence flux.

F_0 dark-adapted minimal fluorescence flux.

F_0' light-adapted minimum fluorescence flux.

F_m dark-adapted maximal fluorescence flux.

F_m' light-adapted maximum fluorescence flux.

F_v dark-adapted variable fluorescence flux defined as the difference between F_m and F_0.

$\mathbf{H}_\theta\,(\theta, \phi)$ Fisher information matrix.

I_α state variable for Taylor series expansion in modified Camacho Rubio model.

I_{α_2} state variable for Taylor series expansion in modified Camacho Rubio model.

I_g growing irradiance [$\mu\mathrm{E/m^2s}$].

I_m measuring light [$\mu\mathrm{E/m^2s}$.

I_{qE} irradiance level at which half of the maximal qE activity is realised [$\mu\mathrm{E/m^2s}$].

k_a kinetic constant of the activation reaction rate in modified Camacho Rubio model [$\mathrm{m^2/\mu E}$].

k_a^{CR} kinetic constant of the activation reaction rate in Camacho Rubio model [$\mathrm{m^2/\mu E}$].

k_a^{EP} kinetic constant of the activation reaction rate in Eilers and Peeters model [$\mathrm{m^2/\mu E}$].

k_c rate constant involved in the photoacclimation process in modified Camacho Rubio model [-].

k_c^{CR} rate constant involved in the photoacclimation process in Camacho Rubio model [$\mathrm{s^{-1}}$].

k_d deactivation reaction rate (photochemical quenching) in modified Camacho Rubio model [$\mathrm{s^{-1}}$].

k_d^{CR} deactivation reaction rate (photochemical quenching) in Camacho Rubio model [s^{-1}].

k_d^{EP} deactivation reaction rate (photochemical quenching) in Eilers and Peeters model [s^{-1}].

$k_{i,0}$ kinetic constant of inhibition reaction rate at low irradiances in modified Camacho Rubio model [$m^2/\mu E$].

$k_{i,1}$ kinetic constant of inhibition reaction rate at high irradiances in modified Camacho Rubio model [$m^2/\mu E$].

k_i^{CR} kinetic constant of inhibition reaction rate in Camacho Rubio model [$m^2/\mu E$].

k_i^{EP} kinetic constant of inhibition reaction rate in Eilers and Peeters model [$m^2/\mu E$].

k_M parameter related to the light variation of the maintenance factor with the light intensity [h^{-1}].

k_p proportionality factor between photochemical quenching and biomass growth rate constant in modified Camacho Rubio model [s/h].

k_p^{CR} proportionality factor between photochemical quenching and biomass growth rate constant in Camacho Rubio model [s/h].

k_p^{EP} proportionality factor between photochemical quenching and biomass growth rate constant in Eilers and Peeters model [s/h].

k_r kinetic constant of the recovery reaction rate in modified Camacho Rubio model [s^{-1}].

k_r^{CR} kinetic constant of the recovery reaction rate in Camacho Rubio model [s^{-1}].

k_r^{EP} kinetic constant of the recovery reaction rate in Eilers and Peeters model [s^{-1}].

K_S^{CR} Michaelis constant for activation reaction rate in Camacho Rubio model [-].

M light dependent maintenance factor in modified Camacho Rubio model [h^{-1}].

$M(\theta)$ generic multi input multi output system.

M^{CR} maintenance factor in Camacho Rubio model [h^{-1}].

M^{EP} maintenance factor in Eilers and Peeters model [h^{-1}].

M_0 maintenance factor in the dark in modified Camacho Rubio model [h^{-1}].

N number of photosynthetic units [$\mu E g_{chl}^{-1}$].

n Hill parameter related to the shape of sigmoid function describing NPQ activity [-].

P photosynthesis rate [mol $O_2/g_{chl}s$].

p probability of excitation energy transfer from two antenna systems in the connected units model.

$\mathbf{Q}$ dynamic sensitivity matrix of experimental measurements.

q maximum quantum yield of photosynthesis [-].

q_L parameter estimating the fraction of PSII reaction centres in open states based on a lake model for the photosynthetic unit [-].

q_{NPQ} NPQ index [-].

q_P parameter estimating the fraction of PSII reaction centres in open states based on a puddle model for the photosynthetic unit [-].

R photosynthesis quantum yield [mol O_2/μE].

R_0 ratio between the inhibition rate in the dark and recovery rate [m^2s/μE].

R_1 ratio between the inhibition rate in the light and recovery rate [m^2s/μE].

R_2 ratio between the activation rate and deactivation rate [m^2s/μE].

r_m^{CR} maximum rate constant for activation reaction rate in Camacho Rubio model [m^2/μE].

S_F scaling factor for fluorescence model (proportional to the chlorophyll content).

$\mathbf{V}_\theta(\theta, \phi)$ variance-covariance matrix.

x_1 fraction of photosynthetic units in resting state according to Eilers and Peeters model.

x_2 fraction of photosynthetic units in activated state according to Eilers and Peeters model.

x_3 fraction of photosynthetic units in inhibited state according to Eilers and Peeters model.

List of Figures

1.1 World primary energy demand by region (source IEA (2011)) 2
1.2 Crude oil price between 1970 and 2011 in US dollars (source WTRG 2014). . . . 3
1.3 (a) Annual global anthropogenic CO_2 emissions ($GtCO_2$/yr) from fossil fuel combustion, cement production and flaring, and forestry and other land use (FOLU), 1750-2011. Cumulative emissions and their uncertainties are shown as bars and whiskers, respectively, on the right-hand side. (b) Total anthropogenic GHG emissions ($GtCO_2$-eq/yr) from economic sectors in 2010. The circle shows the shares of direct GHG emissions (in % of total anthropogenic GHG emissions) from five economic sectors in 2010. The pull-out shows how shares of indirect CO_2 emissions (in % of total anthropogenic GHG emissions) from electricity and heat production are attributed to sectors of final energy use. Source: IPCC (2014) 5
1.4 Predicted energy-related CO_2 emissions by region in 2035 according to IEA (2011) and the change from 2010. 6
1.5 Photosynthetic pathways: light reaction for energy capture and dark reaction to fix carbon dioxide. 9
1.6 The Z-scheme. 10
1.7 Block diagram of a typical biodiesel production process from algal biomass. . . . 11
1.8 Open ponds (from `http://www.seambiotic.com`) 13
1.9 Examples of closed photobioreactors as flat plate and tubular reactors. 13
1.10 Thesis roadmap. 20

2.1 Model development information flow. The block diagram represents the key modelling and experimental activities to be carried for reliable model development. 22
2.2 Geometric representation of the alphabetic design criteria. 27
2.3 Scheme showing the possible fates of light energy absorbed by LHC. Energy absorbed by chlorophyll and resulting in singlet excited chlorophyll ($^1Chl^*$) can be used for photoproduction (green arrow), dissipated as heat (blue arrow), dissipated as fluorescence (red arrow) or lead to reactive oxygen species (ROS) (purple arrow). The photoregulation mechanism is triggered by variations in pH gradient across the thylakoid membrane. The funnel representation has been inspired by Demmig-Adams and Adams (2002). 29
2.4 Representative PAM protocol and outcome. The orange shaded area represents the light irradiance; the red line shows the corresponding fluorescence flux measurements (in volts). 30
2.5 Light induced damages in Nannochloropsis. Cells were treated for one hour of strong illumination (2000 $\mu E/m^2s$) in atmospheric CO_2 and then allowed for recovery overnight at 10 $\mu E/m^2s$. 33

3.1 Information flux of model identification procedure. 38

3.2 In Figure (a) the scheme of EPM is reported; x_1, x_2 and x_3 represent the fraction of Photo-Synthetic Unit (PSU)s in resting, activated and inhibited state, respectively. In Figure (b) the scheme of CRM is reported; a_1, a_2 and a_3 are the number of PSUs in resting, activated and inhibited state, respectively, a_t represents the total number of PSUs. 40
3.3 Measurement (black circles) and predicted values of (a) q and (b) q_L. Red solid lines represent the profiles according to EPM, while the dashed green lines represent the profiles according to CRM. 46
3.4 Growth rate constant predicted by the EPM (solid line) and CRM (dashed line) and experimental values of the growth rate constant (black squares). 46
3.5 Measured NPQ values for different light intensities. 47
3.6 Final values of dynamic sensitivities for mCRM, evaluated at different light intensities. 51
3.7 Final values for dynamic sensitivities for the mCRM model evaluated at different light intensities considering literature values for model parameters. 52
3.8 Measurement (black circles) and predicted values of (a) q and (b) q_L. Blue solid lines represent the profiles predicted by the mCRM model. Red stars in panel (a) represent the experimental data used for model validation. 54
3.9 Biomass concentration profiles at different light intensities predicted by the modified Camacho Rubio model. Black circles represent the experimental measurements 54
3.10 Growth rate constant predicted by the modified Camacho Rubio model (solid line) and experimental values of the growth rate constant (black circles). Red stars represent the experimental values of the growth rate constant for the experiments used in the model validation 55
3.11 Biomass concentration profiles at different light intensities predicted by the modified Camacho Rubio model. Red stars represent the experimental measurements used for the model validation 55
3.12 The upper plot shows the dependence of the maintenance factor (expressed in h^{-1}) from light intensity. The lower plot represent the ratio between the maintenance and the gross growth rate (sum of net growth rate with maintenance) dependence from light intensity. 56

4.1 Schematic representation of the Han model 64
4.2 Circles, triangles and squares represent the measured values of $F_m^{'}$, $F_0^{'}$ and $F^{'}$ respectively. The continuous line refers to the variable Φ_f model, while the dashed line refers to the variable σ model. Grey shaded area represent the light profiles. 67
4.3 Comparison between the predicted and measured fluorescence fluxes $F_m^{'}$ (circles), $F_0^{'}$ (triangles) and $F^{'}$ (squares) in response to various actinic light levels I (grey-shaded area) for the calibration experiment. The dotted and continuous lines are obtained without and with the assumption ϕ_A equals ϕ_C, respectively. . 71
4.4 (a) Comparison between the predicted and measured fluorescence indexes Φ_{PS2} (blue lines, square) and q_L (red lines, triangles) at various actinic light levels I (grey-shaded area). (b) Comparison between the predicted and measured fluorescence index q_{NPQ}. 72

4.5 Large plot: Comparison between the fraction C of damaged RCIIs predicted by the full calibrated model (blue lines) and by the expression (4.15) (blue squares) at various actinic light levels I (grey-shaded area). Small plot: Similar comparison for the variable σ model alternative (red lines and circles). 73
4.6 Comparison between the predicted and measured fluorescence fluxes $F_m^{'}$ (circles), $F_0^{'}$ (triangles) and $F^{'}$ (squares) in response to various actinic light levels I (grey-shaded area) for the validation experiment. 74

5.1 (a) Experimental data of chlorophyll content by Anning et al. (2000) (b) Experimental data of total cross section by Falkowski et al. (1985). 77
5.2 (a) Experimental data by Falkowski et al. (1985) demonstrating the variation of parameter N (b) Experimental data by Sforza et al. (2012) demonstrating the different activity of NPQ for different acclimation states. 77
5.3 Fitting results of the fluorescence fluxes for 3 photoacclimation states. Upper, middle and lower plots correspond to 625, 100 and 15 μE/m^{2}s. 79
5.4 PI curves (a) and evolution of damaged RCIIs (b) for 3 photoacclimation states. Dotted, dashed and continuous lines correspond to 625, 100 and 15 μE/m^{2}s. . . 80
5.5 (a) PI curves for a sample acclimated at 100 μE/m^{2}sand different light protocols: dotted line considers a Type A experiment with an incubation time of 1800 s and dashed line considers a Type A experiment with an incubation time of 3600 s; continuous red line consider a Type B experiment with constant light steps of 300 s; dash-dot line considers Type B experiment with steps of 600 s constant light. Black continuous line represent the steady state PI profile predicted by the model. (b) PI curves obtained for a Type B experiment with 300 s constant light step and different initial conditions for damaged reaction centres. 81
5.6 The squares represent the in silico experiment obtained by model simulation and assuming a Type B experiment with 300 s constant light steps. The continuous line is the predicted profile if a calibration is attempted considering a Type B experiment with 600 s constant light steps; the dotted line consider the predicted profiles if a calibration is attempted considering a Type A experiment with 3600 s incubation time. Table 5.2 and Table 5.3 summarise the results of the two parameter estimations. 84
5.7 The squares represent photosynthesis-irradiance (PI) curve measured by Gentile and Blanch (2001). The continuous line is the predicted profile if a Type A experiment with 1800 s incubation time is assumed. The dotted line is the predicted profile if a Type B experiment with 300 s constant light steps is assumed. Table 5.4 summarise the results of the two parameter estimations. . . 86

6.1 (a) Fluorescence emission kinetics of PSII from dark-adapted acclimated cells were treated with DCMU. The time required for reaching half of the maximum is inversely proportional to the so-called functional PSII antenna size. The fluorescence flux are reported for three different actinic light and normalised by the maximum value (a.u.: arbitrary units). (b) The value of functional antenna size (ASII) is reported for five different light intensities along with the linear fit. The orange envelope represents the confidence region of the linear fitting. . . . 91
6.2 Optimal designed experiment (DOE1) for the simple NPQ model for the two alternative experiment structure considered. 93
6.3 Calibration results for the simple NPQ model and two different values of initial damage C_0: continuous lines consider $C_0 = 0$, dashed lines $C_0 = 0.01$ 94

6.4 (a) Comparison between the predicted and measured fluorescence fluxes $F_m^{'}$ (triangles), $F_0^{'}$ (squares) and $F^{'}$ (circles) in response to a constant light experiment. The grey-shaded area represents the light intensity. (b) Measured value of q_{NPQ} during the recovery phase of experiment RecExp along with predicted values using a single exponential curve or a biexponential curve. . . . 96
6.5 Calibration results for complex NPQ model. Measured fluorescence fluxes $F_m^{'}$ (triangles), $F_0^{'}$ (squares) and $F^{'}$ (circles) are reported along with model predictions. The grey shaded area represents the light intensity. 97
6.6 (a)Sensitivity profile of parameter ξ_F. The red continuous line represent the sensitivity for $F_m^{'}$, the green dotted line represent the sensitivity for $F^{'}$ and the blue dashed line represent the sensitivity for $F_0^{'}$. The protocol to obtain those curves was 60 seconds of strong AL followed by 60 seconds of dark. The actual value of ξ_F used was 0.18 s^{-1}. (b) Dependence of optimal time for the sensitivity of $F_m^{'}$ with respect to the reverse of ξ_F The square represent the optimal sampling time after that the light is switched off, the triangles represent the optimal sampling time after that the light is switched to 2000 μE/m^{2}s. . . . 98
6.7 Optimal designed experiments for complex NPQ model. 99
6.8 Experimental error in long recovery experiment 6.8(b), Ch4Exp 6.8(a), DOE2 6.8(c) and DOE3 6.8(d) along with error profiles predicted by the linear variance model considered in this chapter. The two profiles refers to $F_m^{'}$ and $F^{'}$, each profile is simmetric with respect to the x-axis. 100
6.9 Fluorescence profiles predicted by the model with two calibration experiments (dashed blue profile) and three calibration experiments (continuous red profiles). 103
6.10 Validation of florescence model. The continuous line refers to parameter values reported in Table 6.6 obtained with the calibration set A, the dotted line refers to the parameter values obtained with the calibration set B. 104

A.1 Graphical representation of the puddle (a) and the lake model (b) of the LHC-RC complex. 113

List of Tables

3.1 Parameters of EPM significance and units. 41
3.2 Parameters of CRM significance and units. 42
3.3 Measured biomass concentration profiles at different light intensities[a]. 44
3.4 On the left are reported values of growth rate constant and coefficient of determination R^2 obtained from the linear regression of the growth curves, reported in a semilog scale. On the right are reported measured value of parameter q and its normalised value q_{norm} at different light intensities. 44
3.5 Measured value of parameter q_L at different light intensities. 45
3.6 Estimated values of parameters of the reparametrised mCRM, normalised values (with respect to the initial values) of the parameters, confidence intervals and t-student values for parameters. The reference t-value is equal to 1.67. 53

4.1 Actinic light and fluorescence flux measurements for the calibration experiment. 62
4.2 Expressions of PAM fluorescence fluxes (left part) and fluorescence indexes (right part). 68
4.3 Default values of the constant parameters (left part), and estimated values with confidence intervals of the calibrated parameters (right part). 70

5.1 Parameter values for different photoacclimation states. 78
5.2 Parameter values along with 95% confidence interval and t-values if a PI curve measurement is considered in the calibration set along with the PAM experiment used in Chapter 4. The parameter values are normalised by the values obtained from calibration in Chapter 4. Reference t-value is 1.67. 83
5.3 Parameter estimates, confidence intervals 95%and and t-values 95% considering the PI curve in silico experiment in the calibration set if a wrong light protocol is assumed. Reference t-value is 1.67. Statistically unsatisfactory estimates are indicated by (*). 85
5.4 Parameter estimates, confidence intervals 95%and and t-values 95% considering a PI curve from the literature. The results considered a Type A experiment with 3600 s of incubation time or a Type B experiment with a 300 s step duration. Reference t-value is 1.67. Statistically unsatisfactory estimates are indicated by (*). 86

6.1 Effect of the introduction of σ_{PS2} measurements. The parameters not reported here are not significantly affected by the additional experiment. Normalised values of parameters are reported along with 95% confidence interval and t-values. The reference t-value is 1.65. 92

6.2 The first three column report parameter values estimated using one standard (non-designed) PAM experiment along with 95% confidence interval and t-values. Reference t-value is 1.65. The last three columns report the parameter estimates, confidence intervals and t-values using an optimally designed experiment. Statistically unsatisfactory estimates are indicated by (*). 93
6.3 Normalised values of parameters are reported along with 95% confidence interval and t-values. The reference t-value is 1.65. The normalisation factors are the estimates in the first column of Table 6.2. Statistically unsatisfactory estimates are indicated by (*). 95
6.4 Parameter estimates are reported along with 95% confidence interval and t-values. The reference t-value is 1.65. The calibration set is composed by Ch4Exp, RecExp and σ_{PS2} measurement. 97
6.5 Normalised parameter values, 95% confidence interval and t-values for a calibration set composed by Ch4exp, RecExp one of the optimally designed experiments and σ_{PS2} measurement. The reference t-value is 1.65. 99
6.6 Effect of MBDoE: Parameter estimates are reported along with the normalised values (with respect to the values reported in Table 6.4), 95% confidence interval and t-values. The reference t-value is 1.65. Calibration set A includes the experiments Ch4Exp and DOE3; calibration set B includes the experiments Ch4Exp, RecExp and DOE3. 101

Chapter 1

Introduction

The need for finding economically competitive and renewable energy sources as a substitute to fossil fuels is becoming even more pressing. Microalgae-based processes are considered one of the most promising alternative technology for the production of liquid fuels for transport sector (Hannon et al., 2010; Sheehan et al., 1998; Mata et al., 2010). Their high productivity, their ability to accumulate triacylglycelos (TAGs) under certain stress conditions, and their independence from arable land and fresh water all together put them in a competitive position against traditional oil crops (Chisti, 2007; Williams and Laurens, 2010; Mutanda et al., 2011). Nonetheless, much research is still needed in order for algal-derived biofuel to become a reality, starting with making large-scale microalgae cultivation both economically and energetically sustainable. In fact, the most optimistic previsions are based on crude extrapolation of the productivities obtained in the lab, where conditions differ drastically from those in outdoor culture systems, and no pilot or larger scale demonstration plant has been able to reproduce them as of yet (Quinn et al., 2012; Moody et al., 2014). A better understanding of the underlying biophysical and biochemical processes and their interactions is clearly necessary in order to assess the true potential of microalgae culture systems.

Sunlight provides the energy to support microalgae growth and this energy must therefore be exploited with the highest possible efficiency for optimizing productivity. Algae efficiency in converting solar radiation, however, depends on many environmental factors, including light intensity, temperature, and nutrient availability. Photobioreactor design and operation too can play a major role as the culture concentration, depth and mixing can all affect light use efficiency, and therefore productivity. Optimising microalgae productivity in such a complex environment hinges on our ability to describe, in a quantitative manner, the effect of these various parameters as well as their mutual interactions.

The work presented in this Thesis aims at developing robust and reliable mathematical model to describe the bioprocesses involved in microalgae growth so as to provide for quantitative predictions. Such models prove especially useful in identifying which parameters have the largest impact on productivity, thereby providing a means for enhancing the growth conditions through design and operational changes. They can also provide guidance for genetic

engineering related work by identifying those modifications having the largest potential impact on productivity.

The main objective of this Chapter is to present the motivation of the research effort. First, a general overview of the energy situation is presented with a particular concern to the fossil fuels consumption and their effect on climate change is discussed. Next, the biofuels as a possible substitute of the fossil fuels are presented. After a general introduction the third generation biofuels based on microalgae production are presented. The main advantage and bottlenecks for microalgae-derived biodiesel mass production are discussed. Finally, a review of the modelling effort in the literature is carried out and the main limitation of current models are assessed. The motivation of the work and a roadmap of the Thesis will conclude the Chapter.

1.1 Energy outlook

Recent years have seen a continuously increasing energy demand, regardless of financial crisis. The main driver of the growing trend is the population growth and the supply reliability is forecast to fall. Moreover, the actual energy system is based mainly on non-renewable sources and is recognised to be unsustainable from an environmental point of view.

In 2013 the annual world primary energy consumption was estimated at 12730 million tonnes of oil equivalent (MTOE). Fossil resources in 2013 accounted for 87% of the primary energy consumption with oil (33 %), natural gas (24 %) and coal (30 %) (BP, 2014). According to the International Energy Agency (IEA) global energy demand is set up to grow by 37% by 2040 (IEA, 2014). China, which is currently the largest energy-consuming country, will consolidate its leadership in the next decades. A landmark in the geography of energy consumption is forecast in the early 2030s, when China is expected to become also the largest oil-consuming country, crossing paths with United States. But, by this time, India South-east Asia, Middle East and sub-Saharan Africa are expected to take over as the engines of global energy demand growth (IEA, 2014).

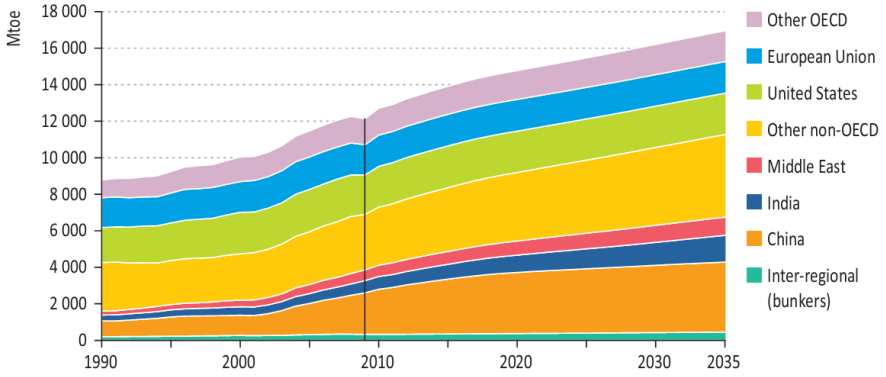


Figure 1.1: *World primary energy demand by region (source IEA (2011))*

The global energy market can be essentially divided into two areas: the power and fuel

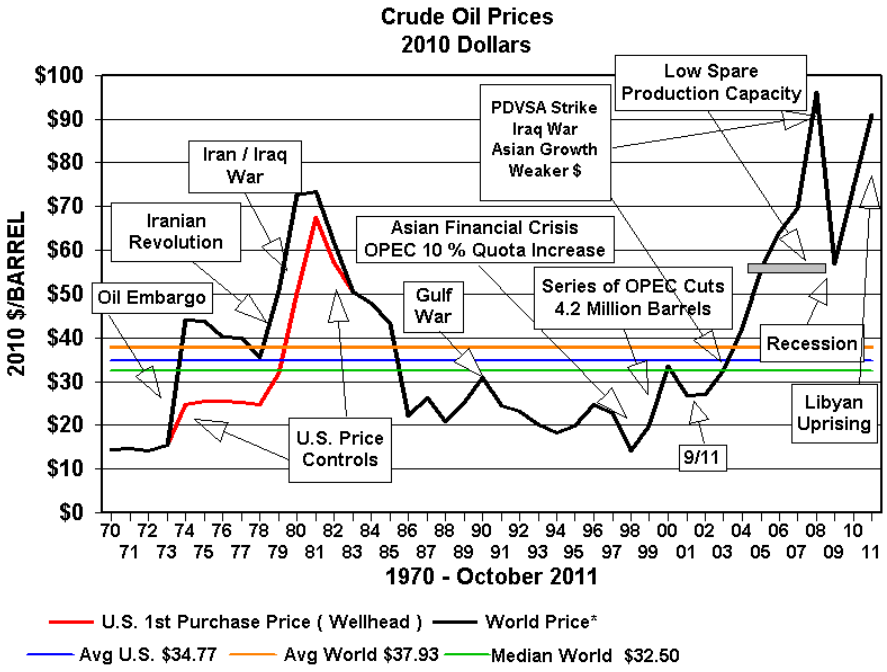


Figure 1.2: *Crude oil price between 1970 and 2011 in US dollars (source WTRG 2014).*

sector. According to IEA, the power sector accounted for the 38% of global primary energy demand in 2009 and is expected to increase up to 42% in 2035 (IEA, 2011). However, the fuel sector will hold the larger share and alternatives to fossil fuel not as obvious, but are advocated to reduce the greenhouse gasses emissions and meet the reduction required from international legislated targets.

Apart from the environmental issues there are also growing concerns about the finiteness of the resources, in particular for the oil production. On the one hand, there are growing evidence that present levels of oil production will probably be insufficient to satisfy the fast growing demand in many countries (among others China and India). On the other hand, concerns have been arising about the so called peak of oil production, i.e. the point in time when the maximum oil productivity is reached, after which the oil production is expected to decrease (Hubbert, 1949). Although the precise date for the peak of oil production is far from being consensual, many experts agree that the world's oil supply cannot expand fast enough to satisfy the growing demand for energy. As a result, they warn, we can expect more price spikes like the ones that have shocked the economy in recent years (Kaufmann, 2011) (see Figure 1.2). At the same time, the petroleum toll on the environment is becoming clearer, especially its huge role in warming the Earth's climate.

1.2 Climate change

Anthropogenic green house gases (GHG) emissions since the pre-industrial era have driven large increases in the atmospheric concentrations of CO_2, CH_4 and N_2O. Between 1750 and 2011,

cumulative anthropogenic CO_2 emissions to the atmosphere were $2040 \pm 310 GtCO_2$. About 40% of these emissions have remained in the atmosphere; the rest was removed from the atmosphere and stored on land (in plants and soils) and in the ocean. The ocean has absorbed about 30% of the emitted anthropogenic CO_2, causing ocean acidification. About half of the anthropogenic CO_2 emissions between 1750 and 2011 have occurred in the last 40 years (Figure 1.3(a)) (IPCC, 2014).

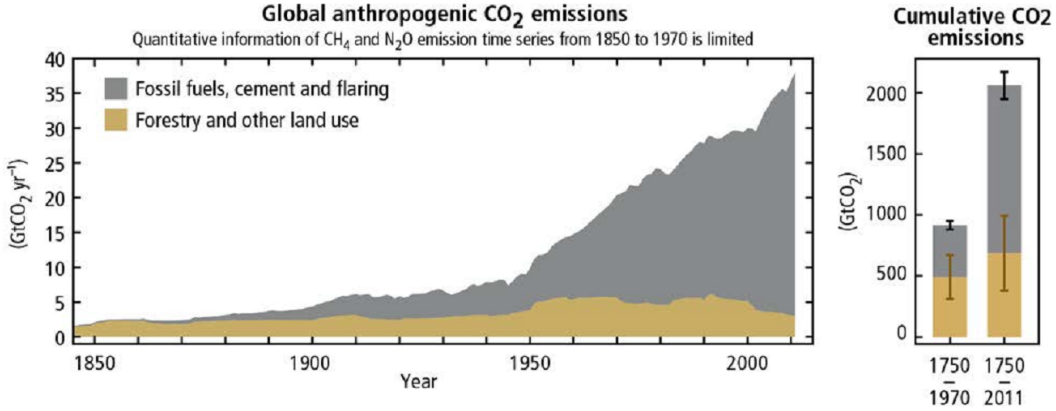
Despite the increased number of climate change mitigation policies, anthropogenic GHG emissions have continued to increase over 1970 to 2010 with larger absolute increases between 2000 and 2010. Anthropogenic greenhouse gas emissions in 2010 reached 49 $GtCO_2$-eq/yr. Emissions of CO_2 from fossil fuel combustion and industrial processes contributed about 78% of the total greenhouse gas emissions increase from 1970 to 2010, with a similar percentage contribution for the increase during the period 2000 to 2010. Total annual anthropogenic GHG emissions have increased by about 10 $GtCO_2$-eq between 2000 and 2010. Since 2000, GHG emissions have been growing in all sectors, except in agriculture, forestry and other land use (AFOLU)(IPCC, 2014).

In 2010, 35% of GHG emissions were released by the energy sector, 24% (net emissions) from AFOLU, 21% by industry, 14% by transport and 6.4 % by the building sector. When emissions from electricity and heat production are attributed to the sectors that use the final energy (i.e. indirect emissions), the shares of the industry and building sectors in global GHG emissions are increased to 31% and 19%, respectively (Figure 1.3(b))(IPCC, 2014).

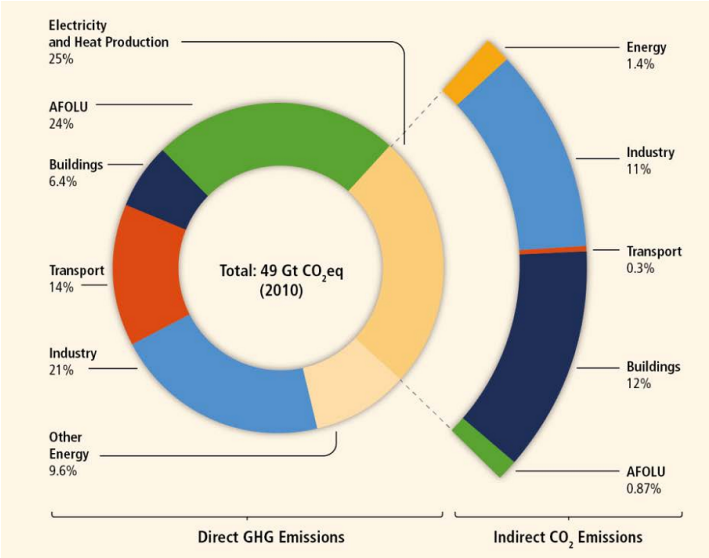
Globally, economic and population growth continued to be the most important drivers of increases in CO_2 emissions from fossil fuel combustion. The contribution of population growth between 2000 and 2010 remained roughly identical to the previous three decades, while the contribution of economic growth has risen sharply. Increased use of coal has reversed the long-standing trend of gradual decarbonization (i.e., reducing the carbon intensity of energy) of the world's energy supply (IPCC, 2014).

According to IEA (2011) the global energy-related CO_2 emissions are predicted to grow by about 6 $GtCO_2$-eq from 2010 and 2035. In Figure 1.4 we can observe that OECD countries are predicted to reduce their emissions (e.g. United States and Europe are expected to reduce their energy-related emission by 15% and 22% respectively). However, the non-OECD countries and in particular China and India more than compensate the reductions achieved by OECD countries and therefore the world total emission will increase.

Previsions of future energy supply show a wide variety of promising approaches to address the energy problem in the power sector. However, the knotty quest refers to the transport sector, relying almost completely on oil and where complexities in establishing new logistics along with the very high requirements in energy content per volume gives little room to present-day substitutes like hydrogen or electricity (IEA, 2011).



(a)



(b)

Figure 1.3: *(a) Annual global anthropogenic CO_2 emissions ($GtCO_2$/yr) from fossil fuel combustion, cement production and flaring, and forestry and other land use (FOLU), 1750-2011. Cumulative emissions and their uncertainties are shown as bars and whiskers, respectively, on the right-hand side. (b) Total anthropogenic GHG emissions ($GtCO_2$-eq/yr) from economic sectors in 2010. The circle shows the shares of direct GHG emissions (in % of total anthropogenic GHG emissions) from five economic sectors in 2010. The pull-out shows how shares of indirect CO_2 emissions (in % of total anthropogenic GHG emissions) from electricity and heat production are attributed to sectors of final energy use. Source: IPCC (2014)*

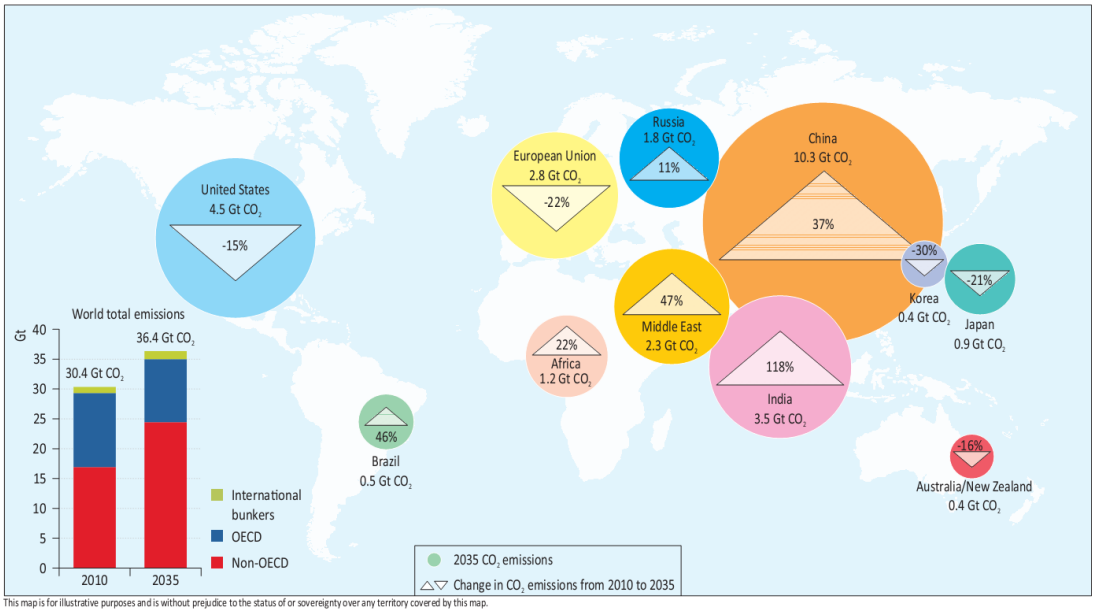


Figure 1.4: *Predicted energy-related* CO_2 *emissions by region in 2035 according to IEA (2011) and the change from 2010.*

1.3 Renewables in transport sector

The transport sector is totally dependent on fossil fuels with a share of about 97% of the supply. Its overall GHG emissions in 2010 has been approximately the 14% of the total GHG emissions (IPCC, 2014) (see Figure 1.3(b)). Global demand for transport appears unlikely to decrease in the foreseeable future; the world energy outlook 2014 (IEA, 2014) projects that increased oil use for transport and petrochemicals will drive demand higher, from 90 million barrels per day (mb/d) in 2013 to 104 mb/d in 2040.

The main drivers in the expansion of transport requirements are again the non-OECD countries and the global increase in the transport sector will be approximately 1.4% per year from 2008 and 2035 (IEA, 2011). To contain the emissions from this sector, policy makers should first and foremost encourage improved vehicle efficiency and increase the share of public transportation and, finally, promote new low-carbon fuels. These include electricity, hydrogen and greater use of biofuels (also as gasoline and diesel fuel blends). One of the main advantages of biofuels with respect to other renewable energy, is the possibility to use them as direct substitute to traditional fossil fuels both for the end users and for the distribution systems.

Biofuels are usually divided into three categories: first generation biofuels are made from traditional crops that are also utilised as food or feed (corn, sugar cane, soybeans; etc.); second generation biofuels are made from non-edible terrestrial biomass such as lignocellulosic material; third generation biofuels aim at not exploiting arable land, microalgae represent their champion.

The main limitation of biofuels utilisation is the high production costs compared to fossil fuels. Moreover, mass production of first generation biofuels have several drawbacks (Sims et al., 2010). One of most common concerns is the *food-for-fuel* dilemma; i.e. they divert food

supply towards the fuel sector, and exploit arable land and fresh water that can be used for food production. The increased pressure on arable land currently used for fuel production could lead to severe food shortage, in particular in developing countries. Palm oil which represent one of the crop with the highest oil producing potential, requires a rainforest-like weather and vast regions, in Brazil, Thailand and Malaysia are currently being cleared at a very fast rate to make room for plantations for biodiesel production (Schenk et al., 2008) resulting in remarkable ecological problems. Another concern is about the high water consumption related to biomass production (Gerbens-Leenes et al., 2009; Bernardi et al., 2012). Second generation biofuels try to overcome the problems related to first generation but the arable land requirement is still present. Another concern related to second generation biofuels is the requirement of a complex and costly collection network that may hinder the economical feasibility (Nigam and Singh, 2011).

Microalgae are considered one of the most appealing solution for the medium term biofuels production (Mata et al., 2010) and extensive research has been carried out to investigate the utilisation of microalgae as energy feedstock. They are considered one of the most promising alternatives to target a newly emerging clean energy market which is predicted to expand rapidly to a value of 500 bn $ by 2050 (Stern et al., 2006).

Several advantages are related to microalgae based biofuels, with respect to traditional crops (Mata et al., 2010). The first one is the extremely high, with respect to terrestrial feedstock, potential oil productivity. However, the absence of industrial scale production system results in high uncertainty in the assessment of the real microalgae oil production potential (the current estimates span two orders of magnitude and the most optimistic projection are based on crude extrapolation of lab-scale results). Recently, Moody et al. (2014) based on a thermal growth model validated with data from outdoor photobioreactors confirm that most of the reported microalgal oil productivities in the literature overestimate the real potential productivity for a large scale microalgae cultivation system, at least in the short-term. However, based on their analysis that can be considered conservative, they conclude that, for example United States and China, the two greatest oil consuming countries, could substitute 30 % of their diesel requirement using only 8 and 5 % of their arable land respectively, or alternatively 11 and 49% of their non-arable land respectively. A second advantage of microalgae is the short harvesting cycle, that allow multiple or continuous harvest, depending on cultivation systems used, in contrast with the traditional crops that are harvested once or twice a year. Moreover, the utilisation of closed photobioreactors could lead to a significant reduction of fresh water requirements.

Finally, another significant aspect is the possibility to enhance with biotechnological techniques the performance of microalgal strains (Radakovits et al., 2010). In fact, it turns out that certain features of microalgae, which have been selected in wild-type species through evolution, can cause a significantly loss of productivity in an artificial environment and are therefore detrimental to large-scale cultivation (Formighieri et al., 2012). One prototypical

example here is the size of the antenna of photosystems, which comprises hundreds of chlorophyll molecules per reaction center in most eukaryotic algae. This large array of pigments maximizing light-harvesting efficiency provide an evolutionary advantage in a natural environment where solar radiation is often limiting for growth and where the ability to be competitive with other photosynthetic organisms is essential to thrive (Kirst and Melis, 2014a). In contrast, such a large antenna is the main cause for non-uniform light distribution in dense microalgae cultures, entailing a significant loss in overall productivity. In fact, the shallow layers are exposed to saturating light and the microalgae present there must activate their protection mechanisms to reduce photo-oxidative damage. Although effective for reducing damages, these mechanisms may end up dissipating up to 80% of the absorbed energy as heat (Barber and Andersson, 1992), strongly reducing light use efficiency in these layers. Then, as most of the light is absorbed in the first few centimeter of the culture, microalgae present in deeper layers become strongly light limited and are therefore at risk of finding themselves below the light compensation point between photosynthesis and respiration. Removing these productivity barriers requires synergies between the development/optimization efforts in photobioreactor design and operation on one hand, and research on genetic modification of microalgae on the other hand. A great deal of research has focused on re-engineering the composition and regulation of the photosynthetic apparatus, so that the modified microalgae can grow faster under dense culture conditions in photobioreactors (Simionato et al., 2013a; Wobbe and Remacle, 2014).

1.4 Photosynthesis fundamentals

Photosynthesis is the key process responsible for the conversion of light energy in chemical energy that is stored in energy rich molecules which can be used for the production of biofuels. A wide range of molecules can be produced by photosynthetic organism: sugar and starch, for bioethanol production, oil, for biodiesel production, biohydrogen, or biomass for biomethane production. In plant and algae the photosynthesis occurs in organells called chloroplasts. Traditionally, photosynthesis is divided into two main reactions sets: the so called “light reactions” of photosynthesis involves the capture of photons and their conversion to energy carriers as ATP and NADPH. At this stage oxygen is produced by the water splitting reaction as a side product. ATP and NADPH are used in the second stage, the so called “dark reactions”, which are independent of light and use the energy stored in the ATP and NADPH molecules to fix the CO_2 producing sugars or other energetic molecules (see Figure 1.5).

In higher plant and microalgae, the light reactions occurs in the thylakoids membrane (see Figure 1.6). The initial step of light reactions is carried out by specialised light harvesting complex proteins, responsible of light capturing, referred as Light-Harvesting Complex of photosystem I (LHCI) and Light-Harvesting Complex of photosystem II (LHCII). These proteins bind the bulk of chlorophyll and carotenoids of the cell and are involved both in light

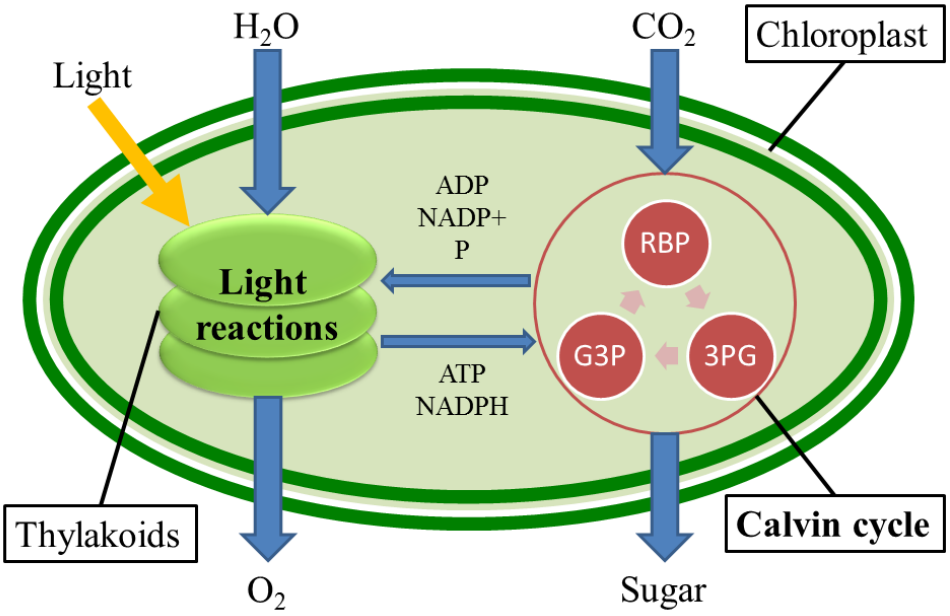


Figure 1.5: *Photosynthetic pathways: light reaction for energy capture and dark reaction to fix carbon dioxide.*

capturing and in the dissipation of excess energy, which otherwise would cause the formation of oxygen reactive species leading to the inhibition of photosystem II (Horton and Ruban, 2005).

The light energy captured by LHCI and LHCII is funnelled to the respective reaction centres of Photosystem I (PSI) and Photosystem II (PSII) via a highly coordinated network of pigments bounded by the LHC. The energy absorbed by LHCII and passed to the PSII reaction centre is used to drive the water splitting reaction, which turns water into electrons, protons, and oxygen. The electrons are shuttled by a series of electron carriers (plastoquinone (PQ), cytochrome b_6f (Cyt b_6f), plastocyanin (PC), PSI) along the photosynthetic electron transport chain following the so-called Z-scheme. The final acceptor is the NPDP+ molecule, which is converted in NADPH, an energetic molecule that will be used in the dark phase of photosynthesis. Simultaneously, a proton gradient is built between the lumen, the inner part of thylakoids, and the stroma, the outer part, by PSII and the PQ/PQH_2[1] cycle. The proton gradient drives the ATP production in the ATP synthase.

The above mentioned process is called linear electron flow. An alternative light energy utilisation is the cyclic electron flow, driven by PSI. The cyclic reaction is similar to the linear reaction but produces only ATP and no NADPH is generated. The cyclic reaction takes place only in PSI: once the electrons are displaced from the photosystem they are passed down through the electron carriers and returns to PSI from where they were emitted, thus the name cyclic electron flow.

The dark reaction of the photosynthesis occurs in the stroma and process is a cyclic reaction, referred as Calvin cycle, which uses ATP and NADPH generated by the light reactions to fix

[1] PQH_2 is the reduced form of PQ

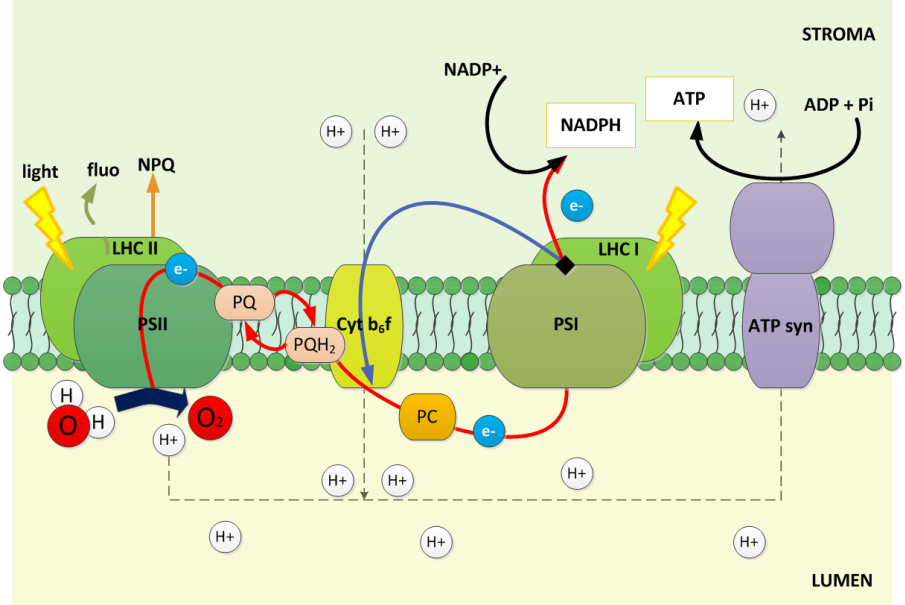


Figure 1.6: *The Z-scheme.*

CO_2. The Calvin cycle can be divided in three main steps. In the first step CO_2 enters the cycle to react with ribulose-1,5-biphosphate (RuBP) to form two molecules of 3-phosphoglycerate (3PG) and is catalysed by ribulose-1,5-biphosphate carboxylase-oxygenase (RUBISCO).The forward reaction is strongly favoured by the negative change in free energy of the process. During the second step, which is dependent upon ATP and NADPH, the 3PG molecules are reduced to form two molecules of glyceraldehyde-3-phosphate (G3P). The third step, consists in a series of reactions that convert a portion of G3P back to RuBP to allow the photosynthetic reduction cycle to continue (Taiz and Zeiger, 2010).

1.5 Third generation biofuels from microalgae

In Figure 1.7 a block diagram of a typical biodiesel production process from microalgae is presented.

The first step of the biodiesel production is of course the biomass cultivation phase. Microalgae can be grown in several types of culture systems that are usually grouped into two main categories: open ponds and closed photobioreactors (described in Section 1.5.2). This stage require nutrients and CO_2 to sustain the microalgae growth and fresh water to compensate the evaporative losses. Biomass cultivation is crucial for the economic and environmental sustainability of the overall biodiesel production process and there is a great research effort to improve both microalgal strains and cultivation systems. The second step is the harvesting of microalgae followed by a dewatering step in order for the biomass to be suitable for lipid extraction. Water separated by biomass needs to be recycled as much as possible to improve sustainability. Oil extraction cost reduction is another challenge that needs to be addressed. The three major alternatives to extract oil from algae are: oil press/expeller, hexane extraction,

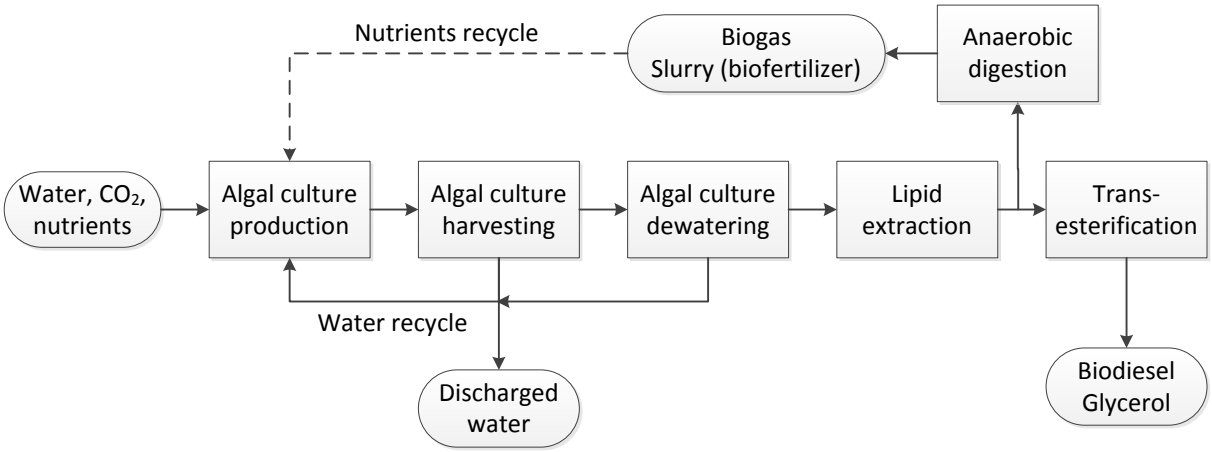


Figure 1.7: *Block diagram of a typical biodiesel production process from algal biomass.*

and supercritical CO_2 fluid extraction (Krichnavaruk et al., 2008). These technologies have all been successfully demonstrated but are relatively expensive, either in terms of equipment need or energy required to extract the oil. Fortunately, all are amenable to engineering improvements (Hannon et al., 2010). Once extracted, the lipids are converted through a transesterification process into biodiesel. At this stage also glycerol is produced as a side product. The solid waste resulting from lipid extraction has to be used in order to enhance the economic and environmental performances of the biodiesel production process. One of the most promising alternative is to perform an anaerobic digestion for the production of biogas (Sialve et al., 2009). The advantage of anaerobic digestion is that the majority of the nutrients are kept in a bacterial slurry that can be sterilised and used for algal fertilizer. Biogas is not currently a high-value commodity, but can help provide energy to operate algae cultivation system.

1.5.1 Environmental challenges and opportunities of microalgae

One of the main potential limiting factors for large scale microalgae-based biofuels is fresh water consumption. Perhaps surprisingly, algae grown in open ponds have water requirements per unit area similar to that of cotton or wheat (but less than that of corn) to replenish the water lost in evaporation (Gerbens-Leenes et al., 2009). However, it is imperative when considering broad deployment of algae, to consider water use to avoid a future water versus fuel debate. In fact, even if microalgae cultivation will occupy non-arable land, water requirement has to be accounted for. Fortunately, many regions have substantial alkaline or saline water reservoirs beneath them, providing a significant source of non-potable water that is suitable for growth of many algal species (Hannon et al., 2010).

Another challenge regards the nutrients supply. In fact, like terrestrial crops microalgae require certain amount of nutrients to grow efficiently. Most of microalgae require to grow phosphorus, nitrogen and potassium (macro-nutrients) as well as micro-nutrients such as iron and sulphur. Most of the fertiliser currently used in terrestrial agriculture to provide nutrients

contain components that are generated from fossil fuels or mined and therefore not renewable (Vance, 2001). A promising alternative to address this issue consists in combining nutrient-rich waste water treatment with algal growth facilities. Extensive works have been conducted to explore the feasibility of using microalgae for wastewater treatment, especially for the removal of nitrogen and phosphorus from effluents (Kong et al., 2010; Chinnasamy et al., 2010; Pittman et al., 2011). Moreover, waste water utilisation will help reducing fresh water requirements

Finally, an important environmental application of microalgae culture is CO_2 mitigation. One of the key advantages, in comparison with terrestrial plants, which typically absorb CO_2 from the atmosphere, is that microalgae can capture CO_2 from high-CO_2 streams such as flue gases and flaring gases (Li et al., 2008).

1.5.2 Microalgae culture systems

As anticipated, the culture system used to cultivate microalgae is a crucial design activity of the overall process. Conventional open ponds and even some closed photobioreactors have already achieved economic viability in the production of high value products such as astaxanthin and nutraceuticals. However, economic margins are much smaller in the biofuel market and therefore it is crucial to optimize the biomass cultivation systems to achieve an economic biodiesel production (Soeder, 1980; Richmond, 2008; Li et al., 2007; Haag, 2007). Recently Jorquera et al. (2010) proposed a comparative life cycle assessment (LCA) analysis for production of biomass using the oil rich microalga Nannochloropsis. The authors compared runaway open ponds, tubular and flat plate photobioreactorss (PBRs) and the net energy ratio (NER) was calculated. The results showed that only runaway open ponds and flat plate PBRs are economically sustainable, with a $NER > 1$; while tubular PBRs have a $NER < 1$, thus underlying the unsustainability of this PBR configuration.

Open ponds

A raceway pond is typically made of a closed loop recirculation channel mixed by a paddle-wheel. In order to prevent sedimentation and enhance mixing, baffles are placed in the flow channel. The first applications of raceway ponds for mass production of microalgae date back to the 1950s (Spolaore et al., 2006). Production of microalgal biomass for biodiesel has been extensively evaluated in raceway ponds in studies sponsored by the United States Department of Energy (Sheehan et al., 1998).

The main advantage of raceways with respect to closed reaction systems is the lower capital and operative cost. However, several issues are related to the raceway operation. First of all, the biomass productivity is lower than a photobioreactor. In addition, carbon dioxide is used less efficiently than in a closed system and they are affected of contamination by unwanted algae and microorganisms that limits the biomass productivity. Finally, open ponds are characterised by high evaporative losses, and the biomass concentration remains low because raceways are

(a) (b)

Figure 1.8: *Open ponds (from* `http://www.seambiotic.com`*)*

poorly mixed and cannot sustain an optically dark zone (Mata et al., 2010).

Photobioreactors

PBRs are defined as reactors in which phototrophs (microbial, algal or plant cells) are grown or used to carry out a photobiological reaction. PBRs are flexible systems that can be optimised according to the biological and physiological characteristics of the algal species being cultivated, allowing to cultivate algal species that cannot be grown in open ponds. PBRs have several advantages over open ponds: they offer better control over culture conditions and growth parameters (pH, temperature, mixing, CO_2 and O_2), prevent evaporation, reduce CO_2 losses, allow to attain higher microalgae densities or cell concentrations, i.e. higher volumetric productivities, offer a safer and more protected environment preventing contamination or minimizing invasion by competing microorganisms (Mata et al., 2010). PBR have been successfully applied for producing large quantities of microalgal biomass (Rubio et al., 1999). However, there are no PBR examples at industrial scale.

(a) (b)

Figure 1.9: *Examples of closed photobioreactors as flat plate and tubular reactors.*

Several PBR designs have been proposed in the literature over the years. Two of the most widely studied configurations are the tubular and the flat panel reactor (Figure 1.9). Despite their advantages, it is not clear if PBRs could have a significant impact in the near future on any product or process that can be attained in large outdoor raceway ponds. PBRs suffer from several drawbacks that need to be considered and solved. Their main limitations include: overheating, bio-fouling, oxygen accumulation, difficulty in scaling up, high capital and operating costs, cell damage by shear stress, and deterioration of material used for the photo-stage. The cost of biomass production in PBRs may be one order of magnitude higher than in open ponds (Mata et al., 2010).

1.6 Models for microalgae growth

A large part of the modelling efforts related to microalgae growth have been developed to account to the growth of phytoplankton in its natural environment. The first phytoplankton model has been developed by Riley (1946) to describe the population of Georges Bank and accounts for light and nutrients effect on the growth (also considering a light exponential attenuation along the depth). It is worth noting that some assumptions that are usually involved in models for phytoplankton in the natural environment are not suitable for an artificial environment. For example, as microalgae evolved over eons to adapt to their environment, it can be assumed that the photosynthetic apparatus is optimally designed to maximise the growth (or more in general the fitness, in an evolutionary sense). When dealing with an artificial environment the algae can be not-optimally adapted as the conditions differs significantly from the natural habitat. The main difference is of course biomass concentration, that in nature is usually very low whereas in an production system must be as high as possible. Another difference is related to the nutrients availability: typically in a natural environment micronutrients, such as iron, can be limiting while in an artificial environment only the macronutrient limitations are of interest, as their optimisation is necessary for the process to be economically sustainable. Finally, the temperature can vary within a wide range on seasonal and also daily basis in open ponds and even more in closed photobioreactors, while in the natural environment temperature variations are much smaller.

A number of contributions regarding the nutrients effect has been proposed in the literature: one of the first models for the study of nutrients uptake was the one of Dugdale (1967); the effect of nutrients on growth has been studied by Droop (1968) and later reformulated by Burmaster (1979). The resulting model is the Droop model, which essentially differs from a classical Monod model by the fact that relates the growth to the intracellular limiting nutrient concentration and not to the dissolved concentration. The Droop model has been extensively validated over the years (Droop, 1983; Bernard and Gouzé, 1999; Vatcheva et al., 2006) and proved to accurately represents the growth in a constant light environment, despite its simplicity. However, the light effect on the growth is not accounted for in the Droop model.

In parallel to nutrient effect, several kinetic models of photosynthesis to represent the impact of light availability have been proposed. The effect of light on microalgae culture systems is intricate. Besides providing the energy need to drive CO_2 fixation, light can be responsible for the production of ROS that lead to a loss of photosynthetic production, a mechanism known as photoinhibition, when the incoming irradiance exceeds a cell capacity to perform its photochemistry. The main target of photoinhibition is PSII and more specifically the D1 protein that disrupts the electron transport chain when damage occurs (Vass, 2012). The continuous repair of damaged D1 proteins allows microalgae to cope with very high light irradiance, making it an effective mechanism to ensure their survival, yet this strategy requires continuous protein synthesis, which incurs a significant energetic penalty and requires large amounts of nutrients (Marshall et al., 2000; Loebl et al., 2010).

Photoinhibition is not the only mechanism influencing light use efficiency. Microalgae have evolved by developing a complex system of regulatory mechanisms to protect their photosynthetic apparatus via dissipating excess energy as heat. This thermal dissipation, the Non Photochemical Quenching (NPQ), is triggered by short-term fluctuations in light and can be activated on a timescale of seconds or minutes. It provides microalgae with an effective protection against photo-oxidative damage, even in the case of drastic variations in light irradiance (Peers et al., 2009). NPQ is activated when light absorption exceeds the light utilization capacity, a situation that leads to the accumulation of protons in the thylakoids lumen. The proteins in the antenna, also called LHC, undergo a conformational change in response to lumenal pH variations, which can result in up to 80% of the total absorbed energy being diverted as heat (Barber and Andersson, 1992). Even though it is crucial for microalgae to protect themselves from high light irradiance in their natural environment, such a massive energy loss drastically hampers biomass productivity in an artificial system. In particular, this constitutes a major drawback in biofuel applications, whereby all energy losses should be minimised. Another key regulatory mechanism in microalgae is photoacclimation, which relates to a cell ability to adjust its pigment/protein content and composition to the growth conditions. These slow regulatory mechanisms, which act on a timescale of days to weeks, are aimed at optimizing the light-harvesting efficiency on account of physiological needs. When exposed to high light irradiance, for instance, microalgae will reduce their light harvesting efficiency via decreasing the chlorophyll content per cell. They will also increase their photosynthetic capacity via enhancing the capacity for oxygen evolution, electron transport and CO_2 consumption (Walters, 2005). In addition, long-term and short-term light responses turn out to be inter-dependent and act synergistically. In the process of acclimating to a higher light irradiance, for instance, a cell may increase its capacity to dissipate energy as heat on the short term as well (Gerotto et al., 2011). In a dynamic environment, microalgae must therefore balance out the needs of efficient light harvesting between sustained growth and sufficient protection against light-induced damage. While different microalgae have achieved this balance through a gradual adaptation over eons, their efficiency is not as high when

cultivated in environments other than their natural habitats, including industrial ponds and photobioreactors.

The development of mathematical models describing photosynthesis as function of light intensity is made particularly arduous by the large number of governing phenomena, acting on multiple time scales from milliseconds to days: photoproduction, occurs in a fraction of a second (Williams and Laurens, 2010); NPQ acts on time scales of seconds or minutes (Eberhard et al., 2008); photoinhibition, on time scales of minutes to hours (Long et al., 1994); photoacclimation, on time scales of hours to days (MacIntyre et al., 2002). The first mathematical model that tries to represent the light effect on photosynthesis is the model by Baly (1935), which describe the photosynthesis as an hyperbolic function of the light intensity. This empirical model has been extended by Vollenweider (1966) to account also for the photoinhibition process. Opposed to the empirical models, mechanistic models are based on the physical and biological laws characterising the phenomena of interest and therefore exhibit a sound predictive capability and can be used to design and optimise the system they represent. Several mechanistic models have been developed over the years and can be divided into two main categories: physiological models and state models.

Physiological models attempt to describe the dynamic behaviour of photosynthetic cells and propose approximations for the actual mechanisms involved in the cells growth. These models may try to represent the optimal allocation of energy and nutrients during cells activities (Ross and Geider, 2009) or to represent a specific metabolic reaction (e.g. Marshall et al. (2000), where the damage and repair cycle of protein D1 is described). Usually these models are extremely detailed and involve a large amount of variables and parameters. The actual identification procedure may be extremely complex (sometimes even impossible) and require numerous, highly specific and costly experiments. For instance, Kroon and Thoms (2006) have recently presented a model of photosynthetic electron transfer in chlorophytes that explicitly considers the entire reaction sequence from Q_A to NADPH. The model includes 31 state variables and 29 rate constants. It can simulate the types of signals generated by fast repetition rate fluorescence and other fluorescence techniques. The main limitation of the model is the numerical complexity. Moreover, despite the high number of variables and parameters some important biological processes, like the photoregulation mechanism, are neglected (and will require additional parameters to be accounted for).

State models are instead based on the concept of PSU and are more instrumental for simulating and optimising industrial cultivation systems. These models are based on the concept of PSU, which is comprised of the antenna and the reaction center, together with the associated apparatus that is activated by a given amount of light energy to produce a certain amount of photoproduct. The name "state models" was coined to reflect that PSUs can be in different states of excitation, and many such models have been proposed over the years, including (but not limited to): Rubio et al. (2003); Eilers and Peeters (1988); García-Camacho et al. (2012); Han (2001); Pahlow (2005); Papadakis et al. (2012); Wu and Merchuk (2001);

Ross et al. (2008). Recently, multiphysics models have started to appear that integrate these state models within CFD simulation in order to represent the effects of light attenuation and mass-transfer limitation (Nauha and Alopaeus, 2013; Hartmann et al., 2014).

Models that try to study the coupled effect of limitations both in nutrients and light have been proposed by Geider et al. (1998), Faugeras et al. (2004) and Pahlow (2005). These models also account for photoacclimation by modelling the chlorophyll content of the cells, which is a variable of the model. More complex models accounting for acclimation are Zonneveld (1998) and Flynn (2001) but being more accurate in the detail of the described mechanisms, they involve more parameters and state variables, which makes their calibration and validation more difficult.

The effect of temperature on growth has been studied with two possible approach: the first one considers light and temperature as independent factors affecting the growth (uncoupled models); the second one aims at describing also the interdependency between the light and the temperature effects (coupled models). The uncoupled models describe the growth as the product of two distinct functions of light and temperature. An example of uncoupled model can be found in Bernard and Rémond (2012): the growth rate at the optimal temperature is multiplied by a function bounded between zero and one (being one the value at the optimal temperature). The curve is an asymmetric bell-shaped curve described in Rosso et al. (1993). Another semi-empirical model of temperature dependence on growth has been proposed by Norberg (2004). On the other hand, coupled models are more accurate as they aim to describe also the interdependencies between the light and temperature effect. For instance, (Duarte, 1995) describe the effect of temperature on growth by expressing some of the parameters of a state model in the form of Arrhenius-like terms. A similar approach has been followed by Dermoun et al. (1992), where the parameters related to photoinhibition are assumed to vary with the temperature. Although coupled models theoretically better represent the impact of temperature than uncoupled models, the limiting step of photosynthesis is not always temperature-dependent (Béchet et al., 2013). In addition, coupled models require a large number of parameters to be fitted experimentally. Therefore, an apparent good fit during validation may only be due to a good adjustment of the set of parameters. This issue, usually referred to as overfitting, can affect the accuracy of the prediction because the model describes noise rather than important trends (Hawkins, 2004).

In conclusion, the available literature models offer a wide range of complexities and consider several limiting factors as independent or interdependent limiting factors. However, we need to recognise that still there is a lack of fundamental knowledge and modelling expertise on a number of key phenomena, which are at the core of microalgae photosynthesis and metabolism. In this Thesis the effect of light intensity on the photosynthetic apparatus will be investigated. The aim is to develop robust and reliable models that could be useful for process scale-up and optimisation. For this reason, the developed models will be based on mechanistic considerations as much as possible. Moreover, a key point of our approach will be to assure the identifiability of

the developed models, i.e. that the optimal parameter values are unique and that these values can be determined in a precise way (Miao et al., 2011). In fact, a non-identifiable model behaves more like an interpolating function and, as such, it may fail to be predictive. Accordingly, the use of such models in assessing design or operational choices is not recommended, even though it may represent some experimental data well.

1.7 Motivation of the work

In view of the above, it is clear that microalgae are among the most promising feedstock to displace, or at least to complement, liquid fuels in the transport sector, where renewable electric power is not (yet) widely applicable (Chisti, 2007). Despite the recognised potential, one of the main issues to address is to bridge the gap between maximal theoretical biomass productivity (or even lab-scale realised productivities) and large-scale realised biomass productivity. The key point is to maximise as much as possible the light utilisation efficiency.

Finding the best compromise between light harvesting efficiency and photoprotection in microalgae cultivation systems calls for a deep understanding of the effect of the key parameters on productivity. Mathematical models can be a great help for this goal. However, as anticipated in the previous section, many of the literature models are oversimplified, or, on the opposite, too detailed resulting in overparametrised models. What we aim to propose is a general model developing approach to guide the model built up. In Chapter 3 we will illustrate such a procedure with respect to literature data by Sforza et al. (2012) coupled with fluorescence measurements. On the one hand, we will demonstrate how literature models, applied out of their development context, may fail to represent the experimental evidences. On the other hand, introducing new processes requires to increase the models complexity and thus the identifiability of the model has to be evaluated. According to the proposed approach both a priori (global if computationally feasible, local otherwise) and a posteriori identifiability will be assessed. Finally, a validation will always be carried out to prove the model predictive capability against data not used for model developing.

The development of a mathematical model using both growth and fluorescence data pointed out another gap of the literature. This gap concerns the utilisation of the chlorophyll fluorescence measurements. In fact, fluorescence measurements are widely used in biological studies and led to important discoveries in the last decades (Baker, 2008), however, they are usually considered only as qualitative indicators of photosynthetic apparatus properties. Traditionally, a number of fluorescence indexes, such as the realized quantum yield of photosynthesis or the NPQ index, have been used for monitoring specific photosynthetic mechanisms, by qualitatively relating these mechanisms to the measured fluorescence fluxes (Roháček and Barták, 1999a). In contrast, little effort has been devoted to quantifying these relations in the form of mathematical models, which would enable accurate predictions of the quantum yield of photosynthesis and in particular of its dynamic response to variable light

conditions. This motivates a more fundamental investigation of chlorophyll fluorescence and a dynamic model incorporating photoproduction, photoregulation and photoinhibition has been developed. In the remaining part of the Thesis (Chapters 4-6), we will show how, the developed model can be used to have quantitative relationship between fluorescence indexes and some key variables of the system, such as the fraction of damaged reaction centres.

Another useful application of the fluorescence model regards the prediction of the photosynthesis rate in terms of oxygen productivity as function of irradiance. In fact, this kind of measurements usually require time consuming and inaccurate experiments, while with our model we can predict the photosynthesis rate, based on fast and and reliable fluorescence measurements. The utilisation of a dynamic model for photosynthesis rate prediction also underline that the usual utilisation of steady state assumption when dealing with photosynthesis rate curves may be inaccurate and lead to severe overestimation of biomass productivity. Finally, a mathematical model of fluorescence can be used to guide the design of information rich experiments, providing a useful tool to further improve and validate the model structure and to enhance the parameter estimation accuracy.

1.8 Thesis roadmap

A basic roadmap is presented in Figure 1.10. The Thesis structure is as follows.
Chapter 2 presents a general model developing framework to achieve a robust and reliable mathematical model and discusses the main mathematical techniques that will be used in this Thesis. Moreover, the chlorophyll fluorescence measurements theoretical basis will be presented along with the current state-of-the-art way of utilising fluorescence measurements.
Chapter 3 aims at developing a mathematical model able to represent some available growth data of Nannochloropsis Salina along with some fluorescence measurements. The model will be developed from existing literature models. An identification analysis will be carried out and reparametrisation techniques will be used to tackle some identification issues. The final model will be shown to correctly represent the experimental data.
Chapter 4 leads to the definition of a semi-mechanistic model of chlorophyll fluorescence, describing the main biological processes that acts in time scales from milliseconds (photoproduction) to hours (photoinihibition), as well as the energy dependent quenching. In particular, the Pulse Amplitude Modulation (PAM) fluorescence experiments will be used for model developing. The model will provide a useful tool to assess the correctness of some hypotheses used in Chapter 3 and will allow the estimation of the dynamic profile of inhibited reaction centres during the fluorescence experiment.
Chapter 5 extends the model presented in Chapter 4 to account for photoacclimation and to be able to predict the photosynthesis rate as a function of light intensity. The dynamic model of photosynthesis rate will unveil some important issues in the usage of experimental photosynthesis-irradiance curves. In addition, the potential advantage of coupling fluorescence

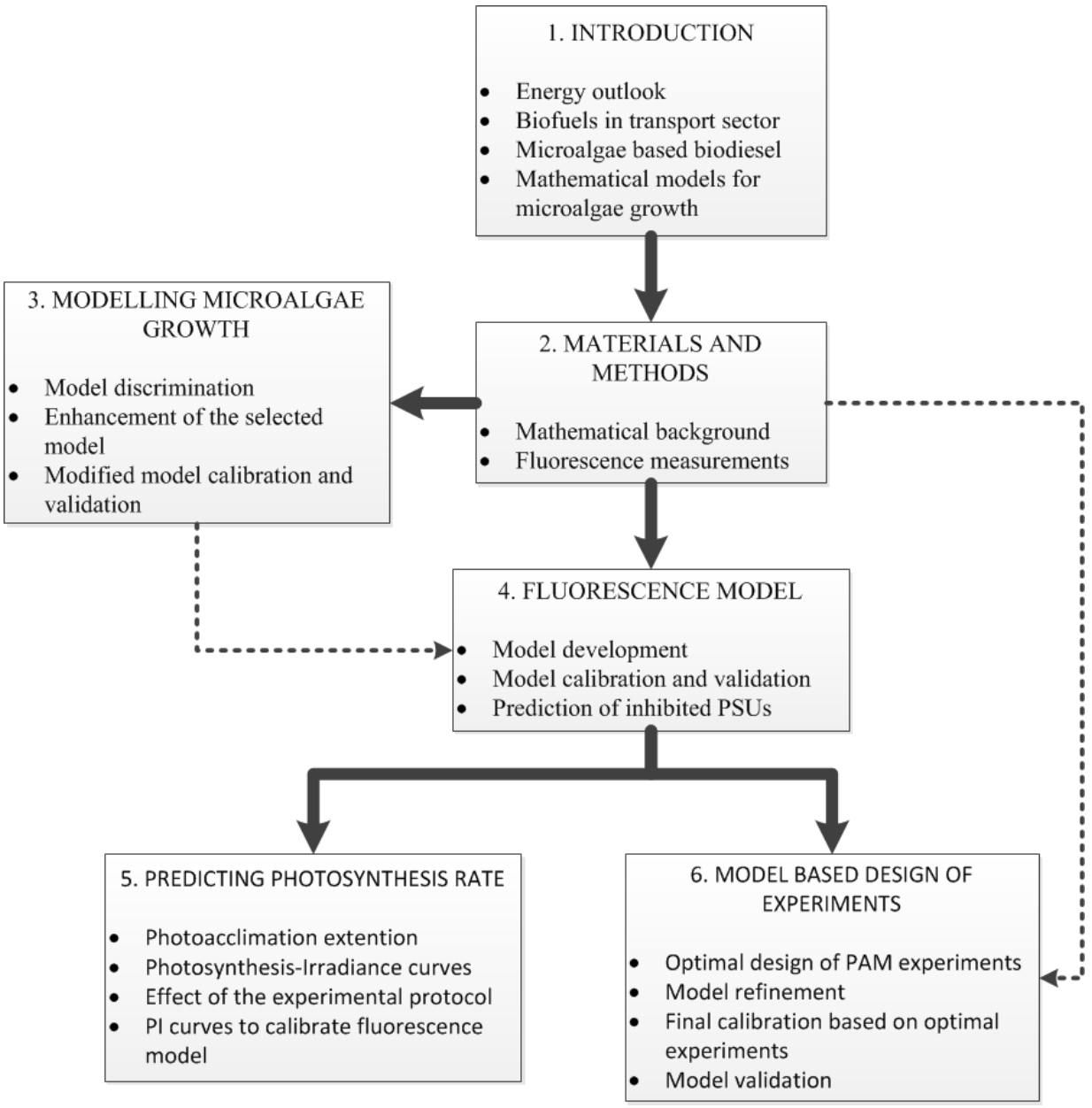


Figure 1.10: *Thesis roadmap.*

measurements with photosynthesis rate measurements will be assessed.
Chapter 6 aims of improving the accuracy of parameter estimation of the fluorescence model presented in Chapter 4. Moreover, Fast Repetition Rate (FRR) fluorescence measurements will be utilised to better characterise the dynamic of the photoproduction.
Chapter 7 summarises the main achievements of the Thesis also suggesting a future roadmap to continue the research initiated through this Thesis.

Chapter 2

Materials and methods

In this chapter[1], are presented the main mathematical tools and experimental techniques used in this Thesis will be introduced. In the fist part of the chapter a general framework for model development is presented and discussed. Next the theoretical basis of model identifiability and experiment design will be explained. The second part of the chapter focuses on the fluorescence measurements, a particular class of measurements widely used in photosynthesis studies and in this Thesis too. The aim of the section is to show how these measurements are currently used in the literature, thus outlining the main limitations related to the standard approach.

2.1 Mathematical background

In this section the main mathematical methods used in this Thesis will be briefly discussed. First, a general procedure for model development will be presented; next, the methods for the global and local identifiability analysis will be introduced. Finally, model based design of experiments (MBDoE) approaches will be discussed.

2.1.1 Developing a first-principles model

The development of a new model is inherently a complex procedure that requires a sound knowledge of the key physical phenomena at play, the ability to formalise this knowledge into a set of algebraic and differential equations, and of course experimental data to support or reject the modelling assumptions. As obvious as this procedure may appear, anyone proposing even the simplest regression model does exactly that. If we start going into more details, however, answers to the following few questions may not look so straightforward anymore: How can we discriminate effectively among several modelling hypotheses? How can we guarantee that a model is identifiable? How can we estimate the model parameters in a statistically meaningful way? How can we deal with uncertainty related to measurement noise, model mismatch and variability in biological responses? And how can we guarantee that a model will be reliable for

[1]Part of this chapter has been published in Meneghesso et al. (submitted)

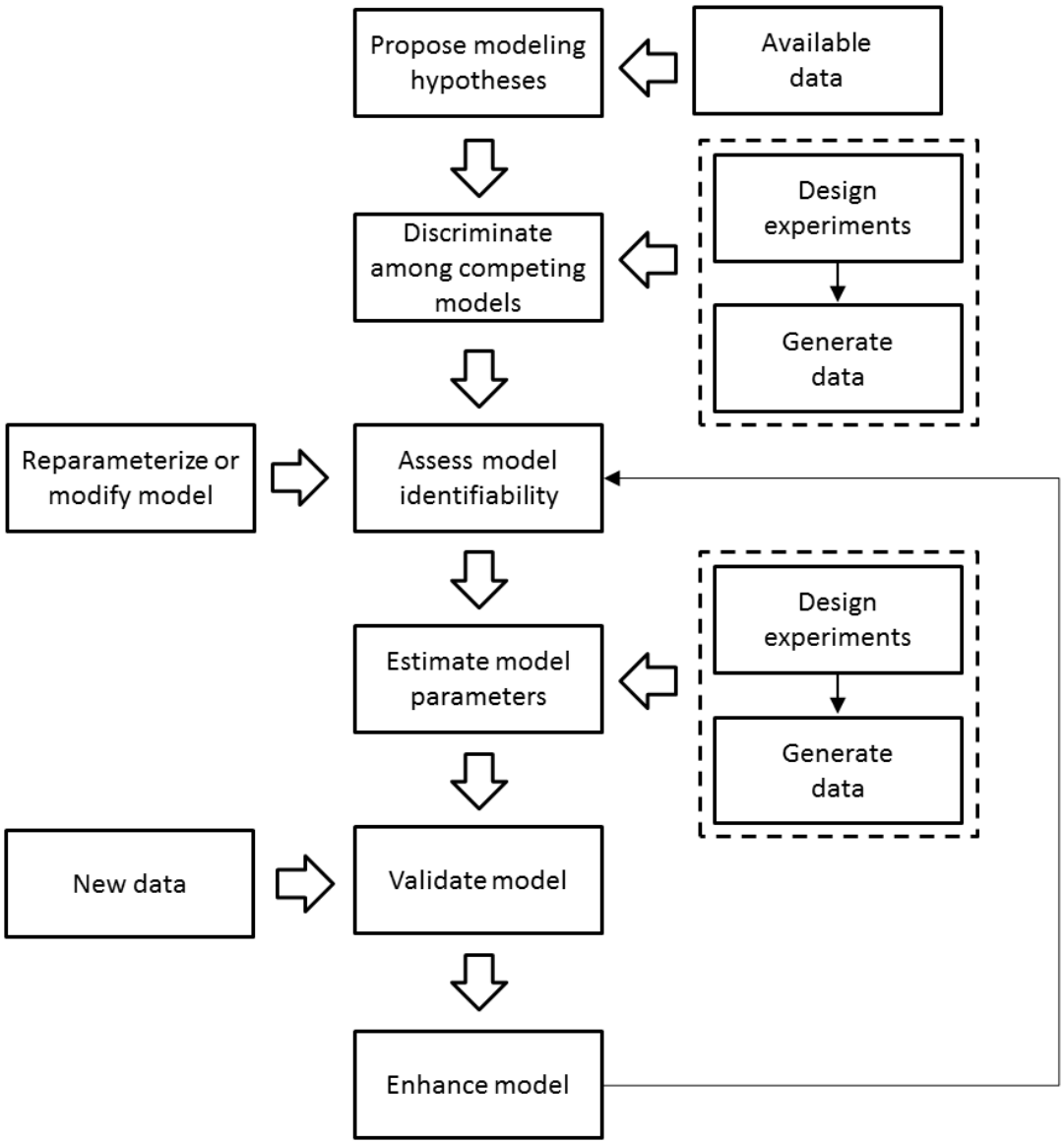


Figure 2.1: *Model development information flow. The block diagram represents the key modelling and experimental activities to be carried for reliable model development.*

the purpose of design or operations optimization?

Figure 2.1 summarises the basic tasks and flow of information that are needed for an effective model development. The preliminary step involves identifying the phenomena that need describing and the fundamental physical principles to represent them. Usually some modelling assumptions are already available, or else preliminary experimental data can be used to decipher correlations among the data and to set up a suitable physical interpretation through a mathematical model. In general, there may be competitive modelling approaches at this stage and the ability to discriminate among them clearly depends on the available experimental data. However, the discrimination procedure can be much more effective, in the sense of minimizing the experimental effort or producing sufficiently informative data, if the experiments are suitably designed. Ideally, input-output data should depend only on the phenomena that

are being investigated and, for this, the interaction between a modelling expert and a biologist is of paramount importance. Moreover, effective discrimination techniques based on advanced methodologies should be considered for the design of such experiments (Chen and Asprey, 2003).

After a suitable candidate model has been selected, the next step involves verifying that the model is identifiable, i.e. that the optimal parameter values are unique and that these values can be determined in a precise way (while still retaining a physically meaning ideally) (Miao et al., 2011). This property is particularly important as multiple parameter combinations may otherwise provide an equally good fit of the experimental data. This would not only lead to a loss of physical significance for the parameters, but a non-identifiable model may turn out to be unreliable when used to simulate or optimise processing/growth conditions differing from the ones used during the calibration. In other words, a non-identifiable model behaves more like an interpolating function and, as such, it may fail to be predictive. Accordingly, the use of such models in assessing design or operational choices is not recommended, even though it may represent some experimental data well. When identifiability issues arise in practice, a first approach to tackle the problem is to re-parameterise the model (Meshkat et al., 2009). If a reparameterisation is still not enough, the model structure should then be modified.

Once a candidate model turns out to be identifiable for the available input-output data, one can proceed with the actual parameter estimation. Note that even though a model is structurally identifiable, measurements noise and other uncertainty effects may still hinder its practical identifiability. MBDoE (Franceschini and Macchietto, 2008) is an effective methodology to address this issue in a systematic way, by determining an experiment that contains the maximum possible information based on a mathematical model. Similar to black-box design of experiments, three consecutive steps are needed to determine the model parameters: (i) design of a new set of experiments based on the current knowledge; (ii) execution of the designed experiment and collection of the new data; and (iii) estimation of the new parameter values and statistical assessment (Asprey and Macchietto, 2000). Multiple Iterations between steps 1-3 lead to a reduction of the uncertain parameter region by progressively adding new experimental information. What is specific to MBDoE compared to black-box DoE here is the use of a system model in step 1 to evaluate the experiment design objective function, as part of a systematic optimization framework based on mathematical programming.

A number of recent MBDoE techniques, including robust MBDoE techniques (Asprey and Macchietto, 2002) and backoff-based MBDoE techniques (Galvanin et al., 2009), have been developed to preserve the experiment design effectiveness despite the presence of significant uncertainty, which can be caused by a large uncertain parameter set, measurement noise, or model mismatch. The integration of backoff-based and online MBDoE techniques (Galvanin et al., 2010) allows for an optimal design of dynamic experiments even when limited prior knowledge is available for the system, with great improvement in terms of design efficiency and flexibility of the overall iterative model development scheme.

Once the final parameter estimation has been performed, different data from the ones used during model calibration should be used to validate the model. Such an interplay between modelling effort and experimental activities allows building-up of model complexity in a controlled manner, keeping in mind the final objective of obtaining a reliable model that can capture the complex multi-scale mechanisms involved in microalgae growth and using this model for optimization and control purposes.

2.1.2 Identifiability analysis

When considering a mathematical model, a crucial aspect to investigate is whether unknown parameters values can be uniquely estimated, based on experimental data. In the literature, several identifiability definitions have been proposed. First of all, identifiability testing can be classified into two categories: a priori identifiability (or structural identifiability testing, based only on the model structure) and a posteriori identifiability (based on collected experimental information).

A priori identifiability (Bellman and Åström, 1970) aims at verifying if, under ideal conditions of noise-free observations and absence of external disturbances, the unknown parameters of a postulated model can be estimated from a designed experiment. Let us consider $M(\theta)$ being a generic multiple-input-multiple-output (MIMO) system that can be described by a nonlinear parametric model:

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

where $\theta \in \Re^{N_\theta}$ is the set of unknown parameters to be estimated, $\mathbf{x}(t) \in \Re^{N_x}$ is the vector of time-dependent state variables, $\mathbf{u}(t) \in \Re^{N_u}$ and $\mathbf{w} \in \Re^{N_w}$ are, respectively, the time-dependent and time-invariant control variables (manipulated inputs), $\hat{\mathbf{y}} \in \Re^{M}$ is the vector of output responses predicted by the model and t is the time. A definition for a priori structural identifiability is given in the following lines.

Definition (structural identifiability): if we denote the equality of the model inputs ($\mathbf{u}(t)$ and $\mathbf{w}$) and outputs ($\hat{\mathbf{y}}(t)$) for two distinct set of parameters $\boldsymbol{\theta}$ and θ^* by $M(\theta) \approx M(\theta^*)$, a parameter $\theta_i \in \boldsymbol{\theta}$ is a priori structurally globally identifiable (SGI) if for almost any θ^*

$$M(\theta) \approx M(\theta^*) \Rightarrow \theta_i = \theta_i^* \tag{2.2}$$

and it is structurally locally identifiable (SLI) if, for almost any θ^*, there exists a neighbourhood $\nu(\theta^*)$ such that 2.2 is still verified (Walter and Lecourtier, 1981).

SLI is a necessary condition to have global identifiability, and a model is said to be SGI if 2.2 is verified for the entire parametric set. A parameter that is not SLI is structurally non-identifiable (SNI) and a model is said to be SNI if any of its parameters is SNI. To test the

identifiability of non-linear parametric models, a local study may be misleading while a global identifiability test should be carried out.

In the literature, several methods to study the structural identifiability of non-linear models have been developed in the last two decades and are nicely reviewed in the work by (Miao et al., 2011). A rather common approach is given by the series expansion method (Pohjanpalo, 1978). It requires that functions representing the model are infinitely differentiable, as the method involve the calculation of arbitrary order derivatives. Several examples of application of this method can be found in literature. However, the series expansion method has a serious drawback: for high dimensional models high order derivatives are necessary and the resulting equations can easily become too complicated to solve.

Ljung and Glad (1994) proposed a method and an explicit algorithm based on differential algebra, demonstrating how the testing of global structural identifiability can be reduced to the question of whether the given model structure can be rearranged as a linear regression. Bellu et al. (2007) developed a specific software tool (named DAISY), based on a new improved differential algebra approach (Saccomani et al., 2003) to test global identifiability of biological and physiological systems. As discussed in Saccomani et al. (2003) a-priori identifiability is a necessary condition (not sufficient) to guarantee successful parameter estimation from real data (a-posteriori identifiability) and, for complex models, the analysis of the model practical identifiability may be the only test that can be carried out, since the a-priori identifiability testing may not be viable because of the computational complexity.

2.1.3 Model based design of experiments

In general an experiment can be described by a series of experimental design variables. We can group all the experimental design variables in one vector, known as design vector, $\boldsymbol{\phi} \in \Re^{N_\phi}$:

$$\boldsymbol{\phi} = [\mathbf{y}_0, \mathbf{u}(t), \mathbf{w}, \mathbf{t}^{sp}, T] \tag{2.3}$$

where $\mathbf{y}_0$ is the set of the initial conditions of the measured variables; $\mathbf{t}_{sp}$ is the vector of N_{sp} sampling times, defining the instants at which the measured variables are sampled; T is the duration of the experiment. In the general case, each of the above mentioned variables can be decided by the experimenter.

Model based experiments design try to define the optimal design vector to achieve a statistically satisfactory parameter estimation. Model based experiment design requires some preliminary information on the model. In particular, we need to have a preliminary estimate of the model parameters θ_0 and a model to describe the experimental error on the measurements. Once these requirements are met the optimal design of experiments consists in varying the vector ϕ in order to optimise a certain metric of the Fisher information matrix (FIM), $\mathbf{H}_\theta(\theta, \phi)$

defined as (Zullo, 1991):

$$\mathbf{H}_\theta(\theta,\phi) = \sum_{r=1}^{N_M}\sum_{s=1}^{N_M} \overline{\sigma}_{r,s}\mathbf{Q}_r^T\mathbf{Q}_s \tag{2.4}$$

where $\overline{\sigma}_{r,s}$ is the r,s-th component of the inverse of the variance-covariance matrix of the experimental measurements $\mathbf{\Sigma}_y$ and $\mathbf{Q}_r$ and $\mathbf{Q}_r$ are the dynamic sensitivity matrix of the r-th and s-th measured variable respectively. $\mathbf{Q}$ is defined as:

$$q_{i,j,k} = \frac{\partial y_i}{\partial \theta_j}|_{t=t_k} \quad \text{with } i = 1,..,N_M;\ j = 1,..,N_\theta;\ k = 1,..,N_{sp} \tag{2.5}$$

If N_{exp} experiments are carried out in series the resulting FIM can be expressed as:

$$\mathbf{H}_\theta(\theta,\phi) = \sum_{j=1}^{N_{exp}}\sum_{r=1}^{N_M}\sum_{s=1}^{N_M} \overline{\sigma}_{r,sj}\mathbf{Q}_{r,j}^T\mathbf{Q}_{s,j} + \mathbf{H}_\theta^0 \tag{2.6}$$

where $\mathbf{H}_\theta^0$ is the initial FIM, that can be neglected if the initial uncertainty on model parameters is very high.

The design problem is usually formulated is an optimisation problem aiming to minimise some metric ψ of variance-covariance matrix, $\mathbf{V}_\theta(\theta,\phi)$, obtained as the inverse of the $\mathbf{H}_\theta(\theta,\phi)$. During the optimisation of the design vector the set of model parameters are kept fixed at the current estimated value $\widehat{\boldsymbol{\theta}}$. The most common optimisation criteria are the so-called alphabetical criteria (Kiefer, 1959):

- D-optimal criterion: the optimisation aims at minimising the determinant of the variance-covariance matrix, $\psi = \det(\mathbf{V}_\theta(\theta,\phi))$;
- E-optimal criterion: the optimisation aims at minimising the larger eigenvalue of the variance-covariance matrix, $\psi = \max_{k=1,..,N_\theta} \lambda_k(\mathbf{V}_\theta(\theta,\phi))$;
- A-optimal criterion: the optimisation aims at minimising the trace of the variance-covariance matrix, $\psi = \mathrm{tr}(\mathbf{V}_\theta(\theta,\phi))$

The geometrical interpretation of the alphabetic design criteria, with reference to a two parameters problem, is illustrated in Figure 2.2.

If we consider linear confidence regions, an A-optimal design aims at decreasing the hyper-rectangular enclosing the confidence ellipsoid. On the other hand, a D-optimal design try to decrease the volume of the confidence ellipsoid. Finally, the E-optimal design aims at minimising the major axis of the confidence ellipsoid, which is defined by the maximum eigenvalue of the variance-covariance matrix. Several other optimisation criteria have been proposed in the literature over the years. Among them we mention the SV-optimal criteria (Galvanin et al., 2007), particularly useful to designing parallel experiments, aims at maximising the information linked to the N_λ largest eigenvalues of $\mathbf{V}_\theta(\theta,\phi)$.

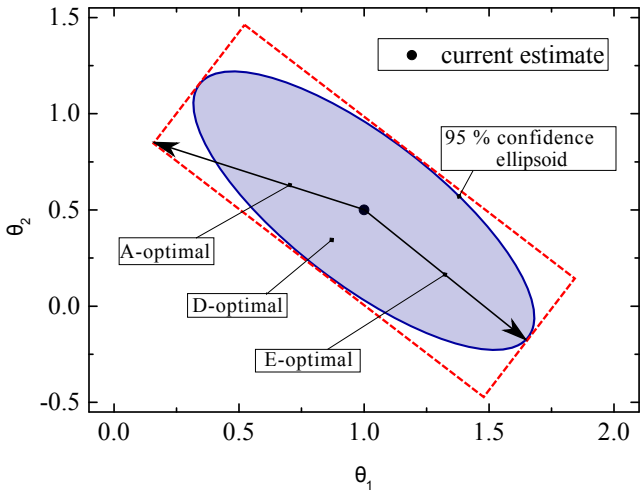


Figure 2.2: *Geometric representation of the alphabetic design criteria.*

2.1.4 Parameter estimation

Once the (designed) experiments have been carried out the parameter of the model should be estimated based on the experimental data. The goal of parameter estimation is dual: the first one is to achieve a statistically sound parameter estimation, providing a precise and accurate[2] parameter estimation; the second one is maximise the ability of the model to represent the experimental data.

Various estimators have been proposed in the literature to achieve the parameter estimation. The simplest one is the least square method, which tries to minimise the sum of squared residuals, a residual being the difference between a measured variable and the fitted value predicted by the model. A slightly more sophisticated estimator is the weighted least square estimator, where the variance-covariance matrix of measurements errors, $\mathbf{\Sigma}_i$, has to be provided for each experimental trial. Accordingly, the objective function to minimise is:

$$\Phi^{WLS}\left(\mathbf{y}, \mathbf{\Sigma}_1, .., \mathbf{\Sigma}_{N_{\text{exp}}}\right) = \sum_{i=1}^{N_{\text{exp}}} \left[\left(\mathbf{y}_i - \widehat{\mathbf{y}}_i\right)^T \mathbf{\Sigma}_i^{-1} \left(\mathbf{y}_i - \widehat{\mathbf{y}}_i\right) \right] \tag{2.7}$$

The least squares methods only provide an estimate of model parameters but do not give a-posteriori statistics concerning the precision of parameter estimation.

For this reason, in this Thesis maximum likelihood method will be used, providing both an estimates of model parameters and the a-posteriori statistics of the parameters. The maximum likelihood method is based on the minimisation of the likelihood function, $L\left(\boldsymbol{\theta}, \mathbf{\Sigma}_1, .., \mathbf{\Sigma}_{N_{\text{exp}}}\right)$, which, when the measurements errors can be considered normally distributed, is expressed as

[2]Note that in case of real experiment the "true" parameter values is obviously unknown and the accuracy of the parameter estimation is difficult to assess

(Bard, 1974):

$$L\left(\boldsymbol{\theta}, \boldsymbol{\Sigma}_1, .., \boldsymbol{\Sigma}_{N_{\text{exp}}}\right) = 2\pi^{N_y/2} \prod_{i=1}^{N_{\text{exp}}} |\boldsymbol{\Sigma}_i|^{-1/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{N_{\text{exp}}}\left[\left(\mathbf{y}_i - \widehat{\mathbf{y}}_i\right)^T \boldsymbol{\Sigma}_i^{-1} \left(\mathbf{y}_i - \widehat{\mathbf{y}}_i\right)\right]\right\} \tag{2.8}$$

The output of the parameter estimation will be the current estimates and some a-posteriori statistics defining the confidence of the estimate. To evaluate the a-posteriori identifiability a widely used method is the t-value test. For every estimated parameter it is calculated:

$$t_i = \frac{\widehat{\theta}_i}{\sqrt{V_{ii}}} \tag{2.9}$$

where V_{ii} is the i-th diagonal term of the parametric variance-covariance matrix. The obtained value is compared with the reference t-value, usually given by a Student t-distribution with $N_{sp} - N_\theta$ degrees of freedom. If the t-value of a given parameter is higher than the reference t-value, the estimate is statistically satisfactory. Very high t-values usually mean that the parameters are estimated with a high confidence.

2.2 Experimental background

In this Thesis data from Nannochloropsis Salina and Nannochloropsis Gaditana will be used to calibrate and validate the developed models. In each chapter the basic informations about the experimental set up used to obtain the data will be specified. The aim of this section is to give a general overview of PAM fluorescence techniques. The overview will be useful to have a clear understanding of the theoretical basis of such a technique and to understand the main limitations related to the state-of-the-art approach of using this kind of data.

2.2.1 Principle of chlorophyll fluorescence

When exposing a photosynthetically active volume to light, a fraction of the light is absorbed by pigment molecules, another fraction is scattered out, and the rest passes through the volume without interaction. In particular, the absorbed photons have four possible fates: they are either captured by the reaction centre of photosystem II (RCII) to drive photosynthesis (photoproduction), dissipated as heat (photoregulation), re-emitted as fluorescence (Müller et al., 2001; Demmig-Adams and Adams, 2002), or lead to the formation of reactive oxygen species (ROS), which are linked to photoinhibition. A nice representation of the energy balance is the so-called funnel scheme reported in Figure 2.3.

The light absorbed by the LHC results in formation of singlet excited chlorophyll (^{1}Chl*). The energy stored in the excited molecule can drive photosynthesis leading to oxygen formation and carbon fixation, or if in excess of what can be used, can be dissipated as heat. Heat dissipation is regulated by the proton gradient between thylakoid membrane, which increases

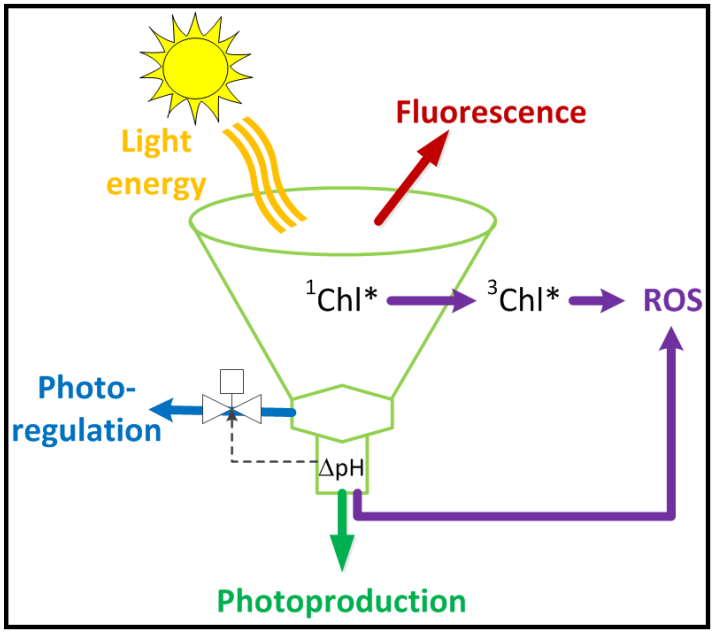


Figure 2.3: *Scheme showing the possible fates of light energy absorbed by LHC. Energy absorbed by chlorophyll and resulting in singlet excited chlorophyll (^{1}Chl*) can be used for photoproduction (green arrow), dissipated as heat (blue arrow), dissipated as fluorescence (red arrow) or lead to ROS (purple arrow). The photoregulation mechanism is triggered by variations in pH gradient across the thylakoid membrane. The funnel representation has been inspired by Demmig-Adams and Adams (2002).*

when the amount of absorbed energy exceed the maximum photosynthesis rate. The rationale of photoregulation is to prevent as much as possible the formation of ROS that can be formed through an over-excitation of chlorophyll molecules into the highly reactive triplet excited state or as side products of photoproduction. In fact, the ROS lead to a net degradation of key photosynthetic proteins, such as the D1 protein of photosystem II (Vass, 2012). In view of the above, it is clear that much information about the photosynthetic processes can be inferred by measuring the fluorescence flux under specific lighting protocols that preferentially activate or inactivate the photoproduction and photoregulation mechanisms.

The funnel scheme described above considers a single reaction centre. When an array of reaction centres are considered a number of configuration can be proposed. In the literature several models have been proposed to describe the LHC-RC complex and are nicely reviewed by Lazr (1999). Among the available models, the most widely used are three: the puddle model, the lake model and the connected units model. The puddle model considers that each reaction centre has his own antenna system. On the other extreme, the lake model considers that a common antenna system is shared by all reaction centres that compete for the excitation energy. The connected units model is intermediate between the puddle and the lake model and considers that each reaction centre has his own antenna system, but, with a certain probability, p, the excitation energy can be transferred from one antenna system to another. It is clear that, if the p is equal to zero the connected units model reduces to the puddle model, on the other hand, while, as p approaches 1 the connected units model approximate the lake model.

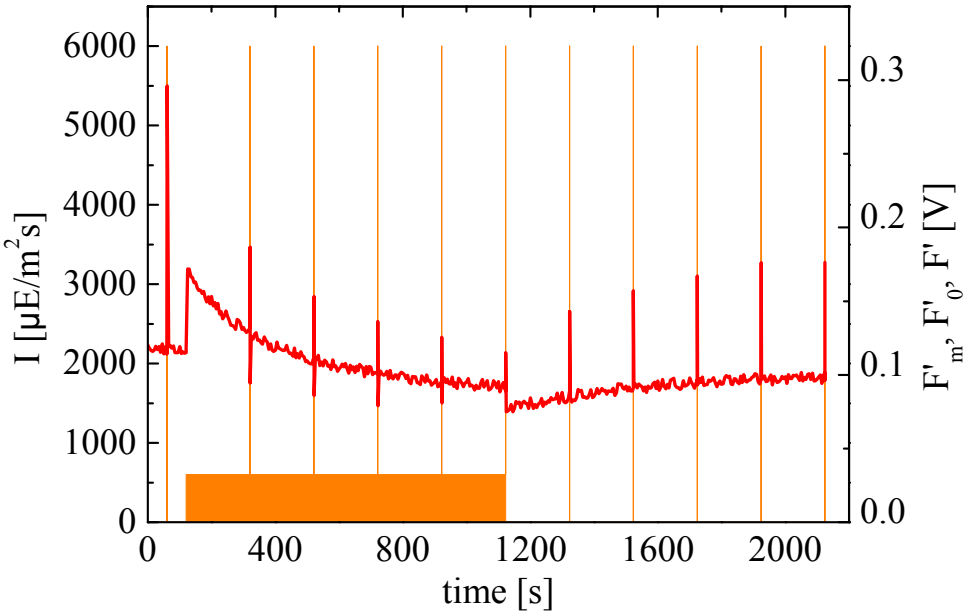


Figure 2.4: *Representative PAM protocol and outcome. The orange shaded area represents the light irradiance; the red line shows the corresponding fluorescence flux measurements (in volts).*

It is clear from extensive research (Butler, 1978; Lavergne and Trissl, 1995; Lazr, 1999; Barber, 2003) that the PSUs do not act as a puddle. A number of elegant measurements have shown that the PSUs are fairly well approximated by the lake model, but are most accurately described by connected units model (Lavergne and Trissl, 1995). However, as suggested by Kramer et al. (2004) connected units models introduce additional free-fitting parameters and according to an Occam's razor approach should be avoided whenever possible. The simpler lake model is usually preferred. In this Thesis the lake model will be assumed to represent the antenna system configuration.

2.2.2 Pulsed amplitude modulation protocols

PAM fluorometry measures the photosynthetic efficiency of photosystem II in a given sample of microalgae, by applying three distinct light: a continuous light (*actinic light*) is used to trigger all the photosynthetic processes, saturating pulses are applied to saturate the photoproduction, and a weak modulated light (*measuring light*) is used to have a fluorescence signal without interfering with the photosynthetic processes, on top of the actinic light used to drive photosynthesis (Roháček and Barták, 1999b). The outcome of a PAM experiment is a record of the fluorescence flux against time, as illustrated in Figure 2.4.

Before conducting a PAM experiment, the microalgae sample is kept is the dark during a sufficient long time in order for (i) all RCIIs to be ready to accept electrons (open state),

and (ii) NPQ to be inactive — the sample is said to be dark-adapted. At the start of the experiment, the measuring light is switched on to a level weak enough (e.g., 0.1 μE/m^{2}s) not to cause significant excitation of the photosynthetic apparatus or trigger NPQ activation — there, the fluorescence detector records the *dark-adapted minimal fluorescence flux*, F_0. Soon after, an intense actinic light pulse is applied (e.g., 6000 μE/m^{2}s), and the detector measures the *dark-adapted maximal fluorescence flux*, F_m. The short duration of the pulse (usually 0.6 s) aims to prevent NPQ activation, while triggering complete excitation of all the RCIIs. Next, the actinic light is switched on at a desired irradiance, so the microalgae progressively transit from dark-adapted to light-adapted state as a result of NPQ activation. During this transition, the detector continuously records the *light-adapted realised fluorescence flux*, F', which is decreasing until NPQ has reached a steady state. Every once in a while, a saturating pulse is applied on top of the actinic light to record the *light-adapted maximal fluorescence flux*, F'_m, and the actinic light is also briefly switched off to record the *light-adapted minimal fluorescence flux*, F'_0. After NPQ has reached its steady state, the actinic light is switched off and recording of the realised, maximal and minimal fluorescence fluxes can continue until the microalgae have reverted back to dark-adapted state. Note that the new dark-adapted state at the end of the experiment may be different from the initial dark-adapted state due to the accumulation of damaged RCIIs (Rees et al., 1990).

2.2.3 Inference of fluorescence protocols: fluorescence indexes

The main fluorescence *indexes*, also commonly referred to as fluorescence parameters in the literature, are expressed as combinations of the characteristic fluxes F_0, F_m, F'_0, F'_m and F' described earlier. By discriminating either between dark- and light-adapted states, or between realised, maximal and minimal excitation states, these indexes allow monitoring of specific photosynthetic mechanisms.

The maximum quantum yield of photosynthesis, q, is given by (Kitajima and Butler, 1975):

$$q = \frac{F_m - F_0}{F_m}\,, \tag{2.10}$$

whereby the difference between F_m and F_0 represents the maximum amount of photons that can be used for photoproduction since NPQ is inactive (dark-adapted). In contrast, the realised quantum yield of photosynthesis, Φ_{PS2}, considers light-adapted states:

$$\Phi_{\mathrm{PS2}} = \frac{F'_m - F'}{F'_m}\,, \tag{2.11}$$

an index also known as the *Genty parameter*, after the researcher who first derived it (Genty et al., 1989). An index related to the Genty parameter, that will be useful for monitoring the

photoinhibition is :

$$\Phi_{\mathrm{L}} = \Phi_{\mathrm{PS2}} \frac{F_0^{'}}{F^{'}} \tag{2.12}$$

Other two useful indexes are:

$$q_P = \frac{F_m^{'} - F^{'}}{F_m^{'} - F_0^{'}}, \quad \text{and} \qquad q_L = q_P \frac{F_0^{'}}{F^{'}}, \tag{2.13}$$

q_P is a widely used parameter to quantify the oxidation state of RCIIs 'in a first approximation, assuming a linear relationship between fluorescence yield and the percentage of open RCII' (Schreiber, 1986). Kramer et al. (2004) demonstrates that parameter q_P is linear with respect to the fraction of open reaction centres only if a puddle model is used to describe the photosynthetic apparatus. Moreover, they introduce the parameter q_L that reflects the fraction of RCII in the open state with respect to the amount of active RCII if a lake model is assumed. In this Thesis the lake model will be assumed to represent the antenna system configuration and therefore q_L will be used to represent the oxidation state of the reaction centres.

Finally, the extent of photoregulation can be monitored through the NPQ index, q_{NPQ}, defined as (Bilger and Björkman, 1990):

$$q_{\mathrm{NPQ}} = \frac{F_m - F_m^{'}}{F_m^{'}}, \tag{2.14}$$

whereby the difference between $F_m^{'}$ and F_m represents the dissipation of energy due to photoregulation.

2.2.4 Utilisation of PAM experiments

Even if the level of understanding of the various fluorescence parameters has significantly increased the past years, little effort has been devoted to the development of dynamic models that associate the operation of the photosynthetic machinery with fluorescence measures. In most literature, the fluorescence indexes are only used to compare different cultures in a qualitative way, but only few contributions try to exploit the quantitative information that can be extracted from those kind of measurements (Wu and Merchuk, 2001).

One prototypical example concerns the photoinihibition process. It is well established in the biological literature the use of fluorescence as a proxy for evaluating photosystems efficiency and the parameter q is commonly employed for evaluation of PSII quantum yield (Maxwell and Johnson, 2000). When cells are exposed to intense light, q decreases as a result of PSII inhibition (Figure 2.5), and when the cells are moved back to dim light, damaged PSII are repaired and PSII quantum yield returns to its normal level within a few hours.

Such a proxy has been extremely helpful in biological investigations of the photoinhibition process, making it possible to isolate the proteins responsible for PSII repair by identifying

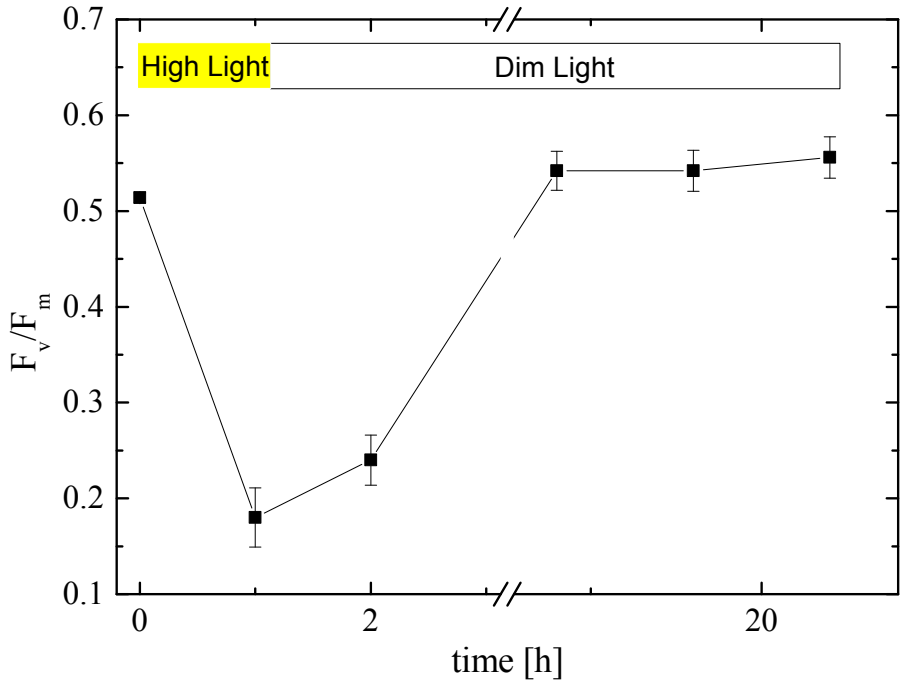


Figure 2.5: *Light induced damages in Nannochloropsis. Cells were treated for one hour of strong illumination (2000* μE/m^{2}s*) in atmospheric* CO_2 *and then allowed for recovery overnight at 10* μE/m^{2}s*.*

mutants from their delayed recovery in q after exposure to a high light irradiance (Park et al., 2007). Although precise, however, q does not quantify the extent of PSII damage. In the work by Wu and Merchuk (2001) a linear relationship is assumed between the fluorescence index q and the fraction of inhibited PSII. This assumption can be acceptable as first approximation but mathematical models describing fluorescence fluxes can help unveiling the actual relationship between the measured parameters and the fraction of inhibited reaction centres.

The only fluorescence index that is quantitatively related to the PSII oxidation state is parameter q_L (Kramer et al., 2004). In particular, q_L is the ratio between open reaction centres and active (i.e. the sum of open and closed) reaction centres. This relationship has been mathematically derived by Kramer et al. (2004) considering the lake model to describe the antenna system configuration.

In Chapter 3 a growth model will be developed starting from existing literature models. Both q and q_L indexes will be considered to discriminate between candidate models and to calibrate and validate the resulting growth model. For q the same hypothesis as in Wu and Merchuk (2001) will be assumed as first approximation. In Chapter 4 a mathematical model of fluorescence will be developed. The model will be calibrated directly on the fluorescence fluxes and will not be necessary to assume any relationship between the fluorescence indexes and the oxidation state of PSII. The developed model will be tested to be consisted with the definition of q_L and will also provide a tool to investigate under which constraints the hypothesis on q made Chapter 3 is valid.

Chapter 3

An identifiable state model to describe light intensity influence on microalgae growth

In this chapter[1] a model describing microalgae growth as function of light intensity will be presented and discussed. The model has been developed starting from literature models and exploiting existing data of biomass growth coupled with fluorescence measurements. Two main topics are dealt with in this chapter: on the methodological side a step by step approach will be presented and applied to guarantee the identifiability of the growth model. Secondly, the utilisation of both biomass concentration (growth curves) and multiple fluorescence measurements will allow shading light on some fundamental phenomena in the correlation between illumination and growth; in particular differently from previous contributions, measurements of the light profile of PSU saturation will be exploited.

The general identification methodology followed to develop the model is first presented. Next, the available experimental data and two candidate models will be introduced and discussed. The successive section is about model discrimination and the enhancement of the selected model. Then, an identifiability analysis and a reparametrisation approach will allow setting up an identifiable model. The performance of the model in describing algal growth will be critically discussed. Some final remarks will conclude the chapter.

3.1 Motivation

Despite the high potential as feedstock for the production of fuels and chemicals, the industrial cultivation of microalgae still exhibits many issues. Yield in microalgae cultivation systems is limited by the solar energy that can be harvested. The availability of reliable models representing key phenomena affecting algae growth may help designing and optimising effective

[1]Part of this chapter has been published in Bernardi et al. (2014) and Meneghesso et al. (submitted)

production systems at an industrial level. In this chapter the complex influence of different light regimes on seawater alga *Nannochloropsis salina* growth is represented by first principles models. Experimental data such as in vivo fluorescence measurements are employed to develop the model. The proposed model allows describing all growth curves and fluorescence data in a reliable way. Most importantly, the model structure is assessed and modified in order to guarantee the model identifiability and the estimation of its parametric set in a robust and reliable way.

Algae growth is affected by several variables such as nutrient availability, temperature, mixing, etc. However, being algae photosynthetic organisms, light is a key variable affecting growth efficiency and kinetics: for this reason the focus of this chapter will be on the representation of its influence on growth. It is worth stressing that the correlation between light intensity and growth is a very complex one, and while low irradiation is limiting, its excess drives to the formation of reactive oxygen species and has inhibitory effect (Li et al., 2009). The choice of investigating light influence is also justified by the fact that the surface of industrial scale photobioreactors or ponds determines the amount of energy they can harvest. In fact, increasing surface do improve overall production, but also energetic and economic costs and thus competitiveness is only reached by increasing light use efficiency.

In this chapter we will consider two literature models: the model by Rubio et al. (2003) and the model by Eilers and Peeters (1988) (initially developed by Eilers and Peeters (1988) and then improved by Wu and Merchuk (2001)). These models are capable of representing the key phenomena of interest in this chapter, and they are reasonably simple so as to limit possible identifiability issues. In the Eilers and Peeters model, the authors assume that if an activated PSU absorbs an additional photon it may become inhibited. For this reason they assume the rate of photoihibition to be proportional to light intensity. It is also assumed that photosynthesis (and by consequence biomass growth) is proportional to the transition between activated state and resting state. Later Wu and Merchuck modify the model of Eilers and Peeters introducing a constant maintenance factor in the description of biomass growth.

In the model by Rubio et al. (2003) both photoinhibition and photoacclimation are considered [2]. Photo-acclimation was at first represented as a steady state process, but more recently extended by the same authors in order to represent its dynamics, together with the effect of NPQ and dark respiration (García-Camacho et al., 2012). This last model is indeed very flexible but the number of model parameters to be estimated is very high and their precise identification may become a long and difficult task, especially if a limited amount of data is available and for this reason it will not be discussed further in this chapter.

For model development and identification we consider experimental data referring to a particular species of microalgae of industrial interest (Nannochloropsis salina) grown in

[2]note that in the original work, and in several others, photoacclimation is instead called photo-adaptation. Photoacclimation is however a more accurate definition: in fact, adaptation refers to the organisms modification during evolution to their environment, thus responses with time scales extremely longer that the ones considered here

non-limiting nutrients conditions and in a flat-plate photobioreactor (Sforza et al., 2012). These datasets were selected because experimental conditions were optimised to minimise all influences on algae growth other than light intensity. In fact, nutrients and CO_2 were provided in excess, but also the photobioreactor light path was minimised to reduce as much as possible light attenuation due to cells shading and scattering. Sforza et al. (2012) demonstrated that this assumption was an acceptable approximation, especially considering that we are interested in representing the exponential growth phase, where nutrients availability is high and cells concentration low. Accordingly, these data represent an accurate description of the influence of light alone on algae growth, minimising the effect of other parameters. It is worth stressing the fact that other phenomena, which also play a major influence on algae productivity in industrial photobioreactors, such as the dark/light cycles due to mixing are not considered in this model.

3.2 Model developing approach

Identifiability is a key issue to guarantee reliability and predictive capability in a model being developed. Figure 3.1 outlines the basic tasks and information flux required to achieve such a target. The preliminary step is to identify the phenomena that need describing and the fundamental mathematical laws that should be implemented to represent them. Here we assume that some modelling assumptions are already available. In other cases, preliminary experimental data may be needed to envisage the correlation among data and to set up a suitable physical interpretation through a mathematical model.

Typically, some available data may be exploited at this stage to choose among competitive modelling approaches through suitable discrimination techniques (Box and Hill, 1967; Stewart et al., 1996). At least in the easiest cases a χ^2 test on experimental data may be sufficient to make the discrimination (Akaike, 1974; Stewart et al., 1998). Ad hoc experiments can also specifically be designed to allow for a more effective and reliable discrimination among different candidates (Alberton et al., 2012; Chen and Asprey, 2003). Once a suitable candidate model has been selected (and a preliminary estimation of its parameter has been carried out), model may need upgrading to improve its capability of representing the phenomena being investigated (new experiments may be needed and possibly designed, and an estimation of all model parameters should be attempted).

Then it is extremely important to verify the model identifiability, i.e. to confirm that the optimal set of parameters values is unique and that their values can be determined in a precise way (and ideally in a physically meaningful way). If some identifiability issues arise, then a first approach to tackle the problem is to reparameterise the model (Meshkat et al., 2009). If this is not enough, then the model structure should be modified.

Once the model is proved to be identifiable, the final parameter estimation can be performed. Note that even when a model is identifiable, measurements noise and other uncertainty effects

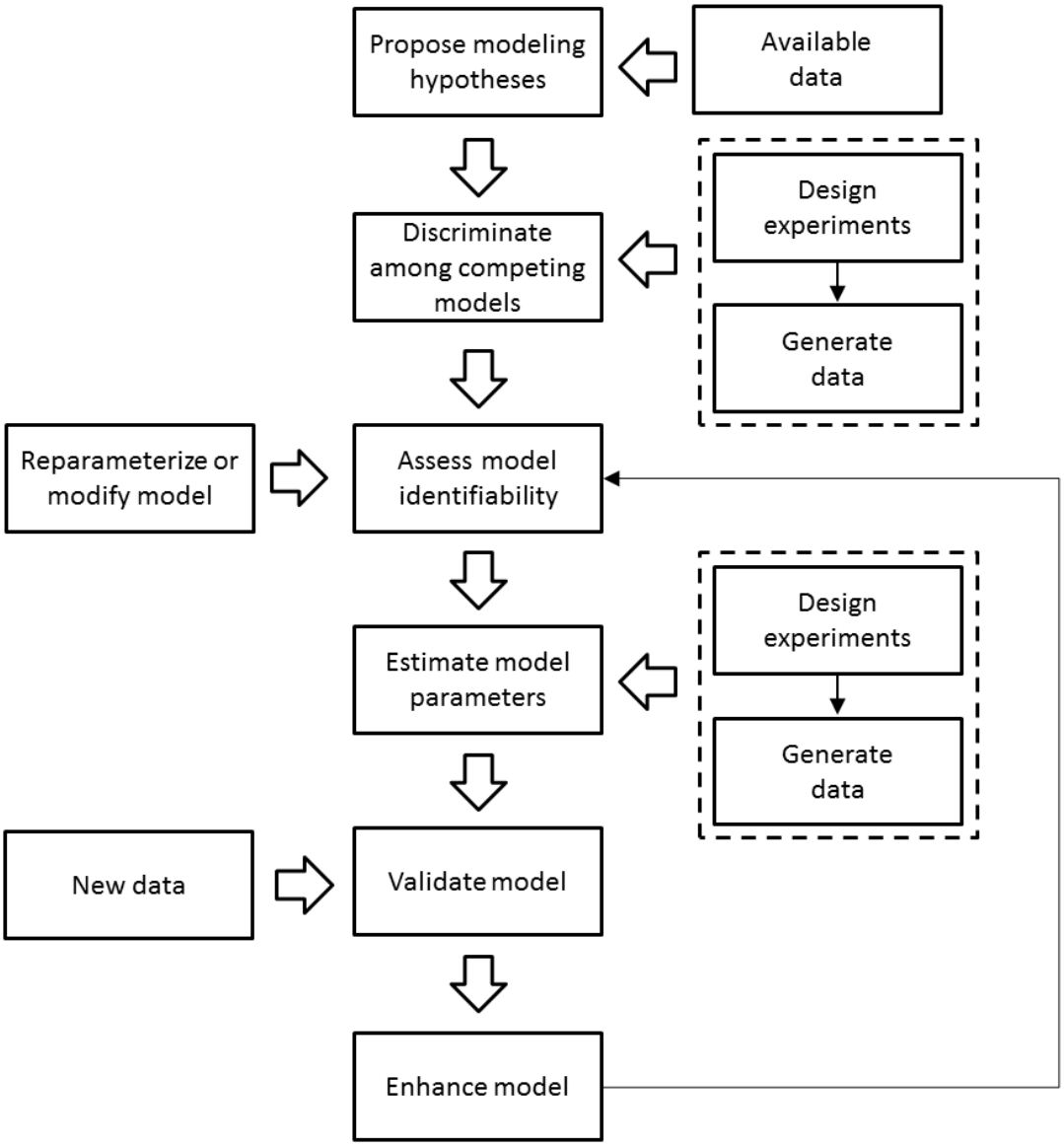


Figure 3.1: *Information flux of model identification procedure.*

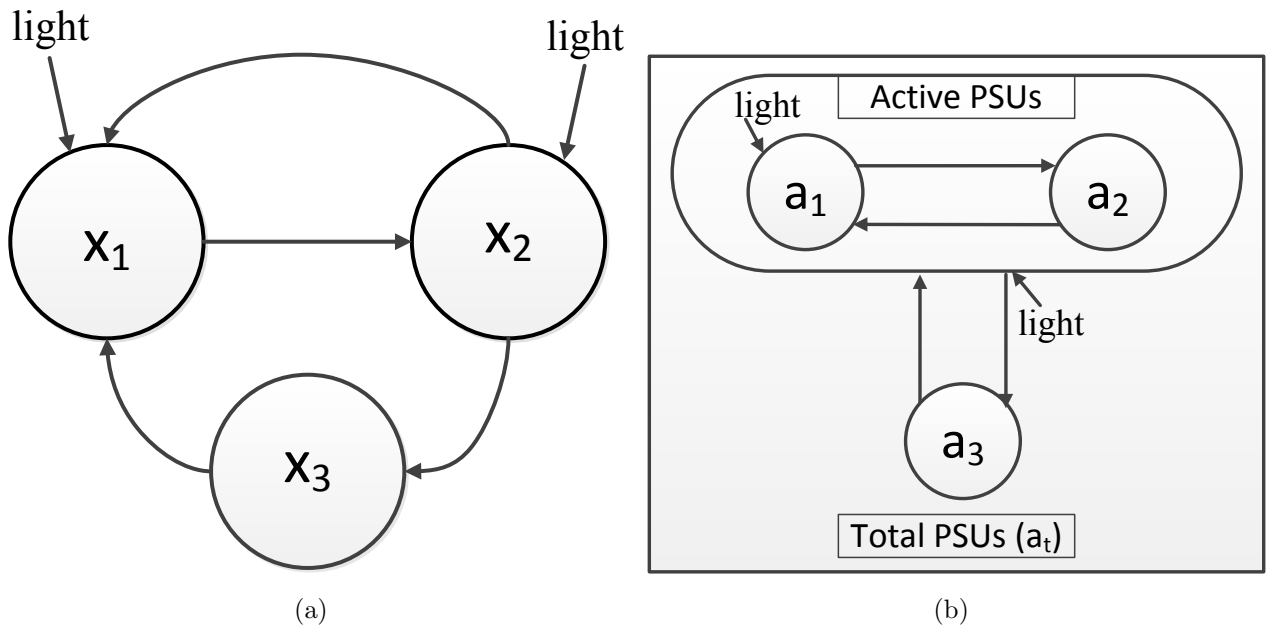
may still hinder its practical identifiability, although properly designed experiments may help tackling the issue (Galvanin et al., 2013). Once the final parameters estimation has been performed, new data, not involved in model calibration, should be used to validate the model. This approach has been applied to the specific case study and is discussed in the following sections.

3.3 Modelling approaches

The main advantage of state models is that they reduce the complexity of photosynthesis into few possible states of the PSUs. This simple structure is also particularly effective in the use of fluorescence measurements, which can be exploited to monitor the PSUs populations at different states. The Eileers and Peeters model in the form proposed by Wu and Merchuk (2001) (afterwards denoted as EPM) and the Camacho Rubio model (Rubio et al. (2003), later called CRM) are two of the simplest models that can describe photosynthetic biomass growth as a function of light intensity. Both models consider that a PSU can assume three different states of excitation: 1) the resting (or open) state, which is the state of PSU before the light energy excites the reaction centre; 2) the activated (or closed) state, that is the state of PSU excited by light energy; 3) the inhibited state, which is the state of PSU damaged by an excess of light energy. Both models do not consider any limitation on nutrients availability or mass transport of nutrients, i.e. only the exponential growth phase is described. CRM considers photo-acclimation too, but without any representation of its dynamics. This is a reasonable assumption also in our case study, since acclimation characteristic time scale (days) is significantly larger than the time scales of the other phenomena being investigated.

In the work of Wu and Merchuk (2001) EPM was used to fit the data of experiments carried out in a thin tubular loop reactor. Part of the reactor was kept in dark to simulate the ordered mixing, and light intensities used for the experiments are 110, 220 and 550 μE/m^{2}s. The measurements used to calibrate the model were both biomass concentration and dark fluorescence measurements. Each experiment was carried out for 48 hours and measurements were taken every 12 hours. In the work of Rubio et al. (2003) Camacho Rubio model (CRM) model was applied to a wider range of light intensities (ranging from 0 to 2000 μE/m^{2}s) and to different light regimes (constant light, flashing light, day-night cycle). Data used by the authors were growth rate constant and P-I curves taken from the literature.

EPM assumes that the number of PSUs is constant with respect to light intensity and accordingly refers to the PSU x_1, x_2 and x_3 to represent the resting, activated and inhibited states, respectively. Conversely, CRM assumes that the number of PSUs is a function of light intensity (indicated as a_t) and the model equations are expressed as a function of the amount of the PSUs in each of the three states (a_1, a_2 and a_3). Note that the photoacclimation effect described by CRM Figures 3.2(a) and 3.2(b) illustrate the two model structures.



(a) (b)

Figure 3.2: *In Figure (a) the scheme of EPM is reported; x_1, x_2 and x_3 represent the fraction of PSUs in resting, activated and inhibited state, respectively. In Figure (b) the scheme of CRM is reported; a_1, a_2 and a_3 are the number of PSUs in resting, activated and inhibited state, respectively, a_t represents the total number of PSUs.*

3.3.1 Eilers Peeters model

In Figure 3.2(a) the three PSU states are represented by circles and the possible state transitions are represented by the arrows. The resting state PSU can capture light energy and transfer it to an activated state. The PSU in the activated state can be damaged by light, or pass down the energy to start the dark phase of photosynthesis (and then return to a resting state). An inhibited PSU can be recovered and then return to the resting state. The reaction rate of the transitions involving the absorption of light (i.e. $x_1 \rightarrow x_2$ and $x_2 \rightarrow x_3$) is assumed to be first order with respect to light intensity. The other two transitions are assumed to be zero order with respect to light intensity. Each transition is assumed to be first order with respect to the PSU fraction involved in the transition. The growth rate constant $\mu^{EP}[\mathrm{h}^{-1}]$ is assumed to be proportional to the state transition from activated to resting state, representing the photochemical reactions. Considering that the growth rate can be negative in the dark or at very low light intensity, a constant maintenance ($M^{EP}[\mathrm{h}^{-1}]$) factor is introduced. The model equations are as follows:

$$\frac{dx_1}{dt} = -k_a^{EP} I x_1 + k_d^{EP} x_2 + k_r^{EP} x_3 \tag{3.1}$$

$$\frac{dx_2}{dt} = k_a^{EP} I x_1 - k_d^{EP} x_2 - k_i^{EP} I x_2 \tag{3.2}$$

$$x_1 + x_2 + x_3 = 1 \tag{3.3}$$

$$\mu^{EP} = k_p^{EP} k_d^{EP} x_2 - M^{EP} \tag{3.4}$$

The set of parameters (whose physical meaning is summarised in Table 3.1 is represented by vector $\widehat{\theta}^{EP} = [k_a^{EP}, k_d^{EP}, k_i^{EP}, k_r^{EP}, k_p^{EP}, M^{EP}]$.

Table 3.1: *Parameters of EPM significance and units.*

Parameter	Significance	Units
k_a^{EP}	Kinetic constant of the activation reaction rate	$m^2/\mu E$
k_d^{EP}	Kinetic constant of the deactivation reaction rate (photochemical quenching)	s^{-1}
k_i^{EP}	Kinetic constant of inhibition reaction rate	$m^2/\mu E$
k_r^{EP}	Kinetic constant of the recovery reaction rate	s^{-1}
k_p^{EP}	Proportionality factor between photochemical quenching and biomass growth rate constant	s/h
M^{EP}	Maintenance factor	h^{-1}

3.3.2 Camacho Rubio model

In CRM, the photoinhibition rate is assumed to be proportional to the sum of resting state and activated PSUs, i.e. the active PSUs represented by $a_1 + a_2$ [PSUs/cells] in Figure 3.2(b). As in EPM, PSU activation reaction is assumed to be first order with respect to light intensity and to the amount of resting state PSUs. However, in CRM the transition from activated to resting state, related to biomass growth, is defined as a Michaelis-Menten kinetic, assuming an enzymatic reaction as limiting step of this process. Also, from an analysis of experimental data, the authors assume that the photo-inhibition reaction rate is first order with respect to the square root of light intensity. As in EPM the recovery of damaged PSUs is assumed to be first order reaction with respect to the number of damaged PSUs. Finally, as anticipated, CRM includes photo-acclimation. The total amount of PSUs (a_t[PSUs/cells]) in CRM is thus assumed to be a hyperbolic decreasing function of light intensity. As for EPM, the growth rate constant is assumed to be proportional to the transition from activated to resting state and a constant maintenance factor ($M^{CR}[h^{-1}]$) is introduced. The model equations are as follows:

$$\frac{da_2}{dt} = -k_a^{CR} I a_1 + \frac{r_m^{CR}}{K_S^{CR} + a_2} a_2 \tag{3.5}$$

$$\frac{da_3}{dt} = k_i^{CR} \sqrt{I}\,(a_1 + a_2) - k_r^{CR} a_3 \tag{3.6}$$

$$a_1 + a_2 + a_3 = a_t \tag{3.7}$$

$$\tag{3.8}$$

$$a_t = \frac{r_m^{CR}}{k_c^{CR} + \frac{k_a^{CR} k_r^{CR}}{k_i^{CR}} \sqrt{I}} \tag{3.9}$$

$$\mu^{CR} = k_p^{CR} \frac{r_m^{CR}}{K_S^{CR} + a_2} a_2 - M^{CR} \tag{3.10}$$

As reported in the literature, it appears that the condition $K_S^{CR} \gg a_2$ is true for the entire range of light intensities considered in the experiments (ranging from 50 to 1000 μE/m^{2}s). Thus, the Michaelis-Menten kinetic may be well approximated by a first order kinetic as in EPM. Accordingly, the reduced set of parameters is represented by vector $\widehat{\theta}^{CR} = [k_a^{CR}, k_d^{CR}, k_i^{CR}, k_r^{CR}, k_p^{CR}, k_c^{CR}, M^{CR}]$ (parameters physical meaning and units are reported in Table 3.2).

Table 3.2: *Parameters of CRM significance and units.*

Parameter	Significance	Units
k_a^{CR}	Kinetic constant of the activation reaction rate	m^2/μE
k_d^{CR}	Kinetic constant of the deactivation reaction rate (photochemical quenching)	s^{-1}
k_i^{CR}	Kinetic constant of inhibition reaction rate	m^2/μE
k_r^{CR}	Kinetic constant of the recovery reaction rate	s^{-1}
k_p^{CR}	Proportionality factor between photochemical quenching and biomass growth rate constant	s/h
k_c^{CR}	Rate constant involved in the photoacclimation process	–
M^{CR}	Maintenance factor	h^{-1}

3.4 Experimental set-up and available data

The aim of this chapter is to describe the growth of microalgae in non-limiting nutrient conditions and according to the hypothesis that the light intensity is constant with respect to the culture time and depth. The fundamental phenomena to be described are: (i) reaction centres oxidation/reduction cycle, to represent the photosynthesis and (ii) the damaging effect of excess light on PSUs (photoinhibition).

Our data refer to algal cultures grown at different light intensities (Sforza et al., 2012). During the experiments the microalgae were acclimated to the light used and to the geometry of the photobioreactor. Each experiment was conducted in parallel at least twice and in two identical photobioreactors, in order to assure its reproducibility. Only the data in the exponential phase of the original growth curves were used for the parameters estimation,

since our model represents only exponential growth and does not consider nutrients limitation. Experimental set up was built to limit as much as possible cells shading, decreasing as much as possible light path and working at low cells concentration. This, together with the presence of nutrients and CO_2 in non-limiting amounts, ensures that growth is dependent only from the light intensity reaching the culture.

Moreover, experimental measurements of fluorescence will be used as additional data for the parameters estimation (Baker, 2008). These data are commonly available for the photosynthetic organisms and have been exploited to estimate photosynthetic efficiency in a large body of experimental literature (reviewed by Maxwell and Johnson (2000)). In this chapter we will use two of the fluorescence parameters introduced in Section 2.2.3: the parameter q (or F_v/F_m) (Barber and Andersson, 1992) and the parameter q_L (Kramer et al., 2004).

The parameter $q = F_v/F_m$ is rather commonly used in similar studies to quantify active PSUs (Wu and Merchuk, 2001). The precise value of F_v/F_m in the case of fully active PSUs is known to be variable between species, since it depends on specific properties such as the antenna size. Here it is set to be equal to 0.65 as this is a value commonly measured in several microalgae (Kolber and Falkowski, 1993) and also in healthy Nannochloropsis cultures, exposed to low light (Sforza et al., 2012; Simionato et al., 2011). In order to have a parameter bounded between 0 and 1 is it possible to define the parameter $q_{norm} = q/q_{max}$, where q_{max} is the value of F_v/F_m in the case of fully active PSUs.

A decrease in this value, in cultures exposed to different light intensities, indicates the presence of photo-inhibited PSUs, as normally experienced in high light conditions. For this reason, it can be used to estimate the content of photo-inhibited PSUs and thus of x_3 and a_3 populations, using the definition of EP and CR models respectively. Note that at this stage we are assuming, as Wu and Merchuk (2001), that the relationship between q and the inhibited PSUs is linear. In the next chapter we will present a more detailed model of fluorescence that will allow us to explain this assumption from a physical point of view in terms of inhibition-related quenching (see 4.3.3).

In order to have a better representation of the oxidation state of the PSUs in illuminated cells, here we also included the fluorescence parameter q_L, which provides a linear estimation of the saturation level of PSU as discussed in detail in the work of Kramer et al. (2004). This means q_L is 0 when all active PSUs are open, while it increases to 1 when all PSUs are closed and photosynthesis is saturated. Parameter q_L was measured for 21 different light intensities with a PAM fluorometer. Thus, q_L can be exploited as an estimation of the relative ratio of x_1 and x_2 populations (a_1 and a_2 according to CRM). The complete set of experimental measurements used in this chapter are reported in the following.

Table 3.3 reports the growth data at different light intensities. Each datum is the mean of two experiments conducted in parallel in two identical reactors. Table 3.4 contains the values of the growth rate constant as obtained from the linear regression of the growth curves reported in a semi-log scale; the coefficient of determination R^2 is also included. From

every biomass culture, a sample was taken during the exponential growth phase and the value of q was measured: in Table 3.4 the available measurements are reported. Finally, the second fluorescence parameter q_L was measured for 21 different light intensities using a PAM fluorometer. The data points obtained by the PAM fluorometer are reported in Table 3.5.

Table 3.3: *Measured biomass concentration profiles at different light intensities*[a].

I [μE/m^{2}s]	50	120	150	250	350	550	750	1000
t [h]	concentration [g/L]							
0	9.93E-02	1.81E-01	1.42E-01	5.49E-02	1.30E-01	1.80E-01	1.21E-01	3.02E-01
24	1.44E-01	3.07E-01	2.40E-01	1.10E-01	1.69E-01		1.61E-01	4.12E-01
36			3.60E-01					
48	1.57E-01	4.69E-01	4.68E-01	2.21E-01	2.02E-01	4.10E-01	1.86E-01	
60			5.76E-01					
72	2.07E-01	8.01E-01		3.32E-01	4.06E-01	6.19E-01	3.50E-01	4.56E-01
84			1.07E+00					
96		1.43E+00		5.57E-01	5.17E-01	7.87E-01	4.80E-01	5.70E-01
120			1.78E+00	6.92E-01		9.50E-01	6.30E-01	1.22E+00
144	6.31E-01			8.32E-01		1.61E+00	1.50E+00	

[a] Experiments carried out at 350 and 750 μE/m^{2}shave been used for mCRM model validation.

Table 3.4: *On the left are reported values of growth rate constant and coefficient of determination R^2 obtained from the linear regression of the growth curves, reported in a semilog scale. On the right are reported measured value of parameter q and its normalised value q_{norm} at different light intensities.*

I μE/m^{2}s	μ [h^{-1}]	R^2
45	0.0126	0.976
120	0.0212	0.998
150	0.0215	0.987
250	0.0190	0.954
350	0.0152	0.953
550	0.0144	0.984
750	0.0145	0.975
1000	0.0107	0.875

I μE/m^{2}s	q	q_{norm}
0	0.650	1.00
45	0.645	0.992
120	0.611	0.940
150	0.592	0.911
250	0.533	0.820
350	0.509	0.783
550	0.503	0.773
750	0.516	0.794
1000	0.499	0.769

3.5 Model discrimination and preliminary parameter estimation

The first objective is to discriminate between the two alternative models so as to select the most suitable one to describe our system. Parameters estimations were performed for both models, based on the entire set of experimental data. The different performance is quantitatively summarised by the χ^2 test (in the case of EPM, we have χ^2= 890.1 ; whereas for CRM, χ^2= 301.6).

Figure 3.3 shows the behaviour of the two models in representing fluorescence experimental profiles. Although both fitting are quite unsatisfactory, CRM clearly outperforms EPM. In

Table 3.5: *Measured value of parameter q_L at different light intensities.*

I μE/m^{2}s	q_L
0	0.65
14	0.57
21	0.55
30	0.54
45	0.53
61	0.52
78	0.52
103	0.50
134	0.46
174	0.44
224	0.40
281	0.37
347	0.33
438	0.28
539	0.23
668	0.20
833	0.17
1036	0.14
1295	0.11
1602	0.087
1960	0.059

fact, Figure 3.3(a) shows that although EPM provides a slightly better fit for the profile of q at low light intensities, only CRM is capable of representing the trend at a higher intensity. Figure 3.3(b), however, shows that, although EPM performance is still the worst one, both models cannot properly represent the oxidative state of PSU in light adapted cells, consistent with the fact that these kind of measurements were not considered in building such models. This means that both models are not accurate in estimating the PSU oxidative state.

Also in the predictions of the growth rate constant, reported in Figure 3.4, CRM outperforms EPM. In fact, EPM does not predict the correct value of light intensity at which the maximum growth rate is reached and underestimates the growth rate constant at low light. CRM correctly predicts the optimal light intensity but it overestimates the growth rate constant at low light.

No additional experiments are needed for discrimination purposes and CRM is then selected, although the experimental data demonstrate that further improvements are required to improve the fitting in the region between 150 and 700 μE/m^{2}s with respect to the PSU oxidation state. It is worth underlining the main differences between EPM and CRM in order to better understand the model modifications introduced in the following section. The first difference between the two model is the presence in CRM of a photo-acclimation estimation in terms of number of PSUs per cell. This suggests that the inclusion of this response is absolutely necessary for an accurate description of the photosynthetic performances and is consistent with its biological relevance, demonstrated by the fact that acclimation responses are conserved in all photosynthetic organisms. The second difference is the order of the photoinhibition reaction rate with respect to the light intensity. In EPM we have a first order reaction while in CRM

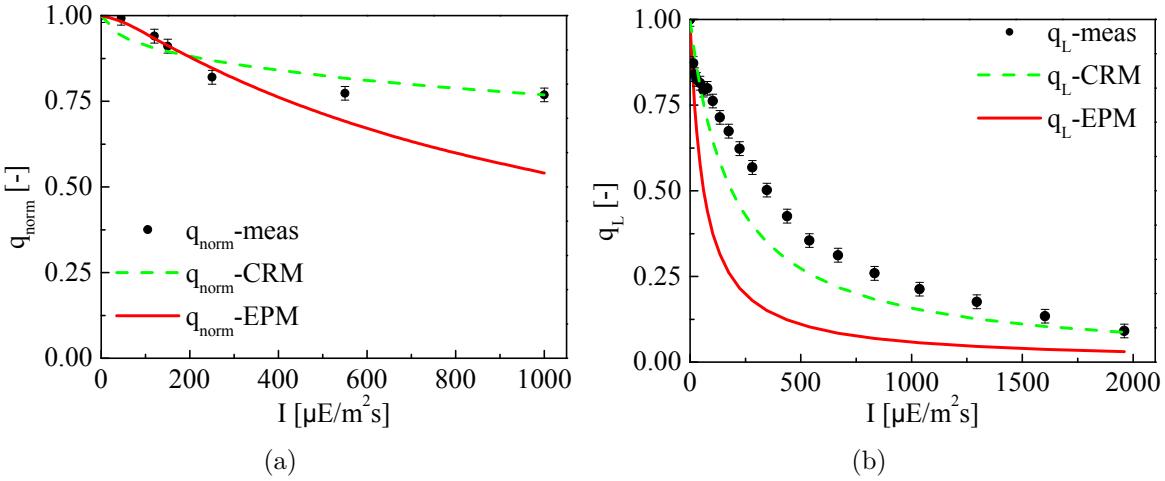


Figure 3.3: *Measurement (black circles) and predicted values of (a) q and (b) q_L. Red solid lines represent the profiles according to EPM, while the dashed green lines represent the profiles according to CRM.*

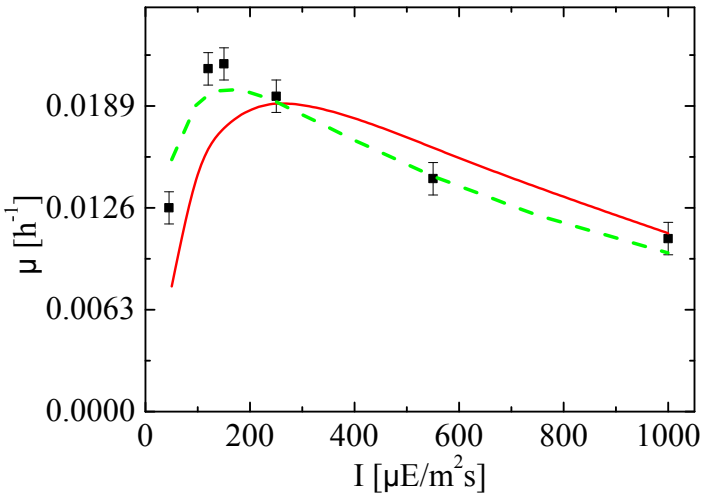


Figure 3.4: *Growth rate constant predicted by the EPM (solid line) and CRM (dashed line) and experimental values of the growth rate constant (black squares).*

the order is 0.5. Although the value of 0.5 does not have a clear justification from a biological point of view it can be interpreted as an effect of photoprotective mechanism, such as NPQ.

3.5.1 Enhancing CRM

To improve the model, a more detailed description of some fundamental biological phenomena needs introducing. According to several works in the literature, PSII photoinhibition occurs at all light intensities (Nixon et al., 2010; Aro et al., 1993; Miyao, 1994). Therefore, we assumed that photoinhibition does not depend on the number of active PSUs, but is simply related to light intensity. Above I_{cr} (the light intensity where photosynthesis is saturated) a second process (photoprotection) is activated: in these conditions a fraction of the energy absorbed does not lead either to photochemistry or to PSU damage, but is simply dissipated (e.g., because

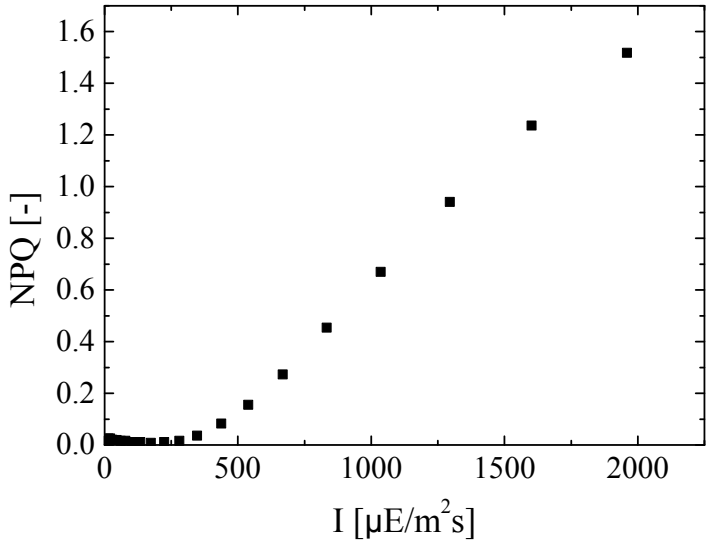


Figure 3.5: *Measured NPQ values for different light intensities.*

of NPQ). Accordingly, Equation 3.6 has been modified to represent the different behaviour above and below I_{cr}. Furthermore, in Equation 3.6, the reaction order with respect to the light intensity is assumed to be 0.5: since there is no clear physical reason for setting such a value, here we decided to increase the model flexibility and its capability of incorporating all energy dissipation phenomena, by making the reaction order a parameter (α) to be estimated. As mentioned above, photoprotection mechanisms are activated when photochemical reactions are saturated, and accordingly parameter I_{cr} plays a key role to represent this behaviour. The value of critical light intensity when photosynthesis is saturated has been fixed to 150 μE/m^{2}s, which is the light intensity in response to which Nannochloropsis growth is maximal, and the limit over which the growth rate is not linearly dependent on light anymore (Sforza et al., 2012). This choice is well consistent with the observation of NPQ dependence from light intensity observed in Nannochloropsis cells, which shows activation only over this limit (Figure 3.5). This parameter thus depends on both light intensity and the number of active PSUs.

Also Equation 3.9 representing photoacclimation is modified to allow for a higher flexibility: the hyperbolic form is retained, but the light exponent becomes an additional parameter (α_2) to be estimated (see later on, Equation 3.14). Finally, the representation of maintenance factor is modified, too. In CRM the maintenance factor is treated as a constant (in fact, this is a typical assumption). However, maintenance factor should vary with light intensity to account for the metabolic cost of repairing damaged PSUs (Jansen et al., 1999). In order to do this, the easiest way is to express the maintenance factor as a linear function of the damaged PSUs fraction. Since the fluorescence measurements return the number of active PSUs, the variable term of maintenance factor was related to the difference between the maximum value of fluorescence (q_{max}) and the current value of fluorescence (q). Finally the fluorescence measurements q and q_L are related to the oxidation state of the PSUs, as discussed in Section 3.4.

The modified Camacho Rubio model (mCRM) is thus constituted by the following set of

equations:

$$\frac{da_2}{dt} = -k_a I a_1 + k_d a_2 \tag{3.11}$$

$$\frac{da_3}{dt} = \begin{cases} k_{i,0} I a_t - k_r a_3 & \text{if } I \leq I_{cr} \\ k_{i,0} I_{cr} a_t + k_{i,1} \left(I - I_{cr}\right)^{\alpha} - k_r a_3 & \text{otherwise} \end{cases} \tag{3.12}$$

$$a_1 + a_2 + a_3 = a_t \tag{3.13}$$

$$a_t = \frac{1}{k_c + I^{\alpha_2}} \tag{3.14}$$

$$\mu = k_p k_d a_2 - M \tag{3.15}$$

$$M = M_0 + k_M \left(q_{max} - q\right) \tag{3.16}$$

$$q = q_{max} \frac{a_1 + a_2}{a_t} \tag{3.17}$$

$$q_L = \frac{a_1}{a_1 + a_2} \tag{3.18}$$

where the new vector of model parameters is $\widehat{\theta} = [k_a, k_d, k_{i,0}, k_{i,1}, k_r, k_p, k_c, k_M, M_0, \alpha, \alpha_2]$.

3.6 Identifiability analysis

In order to be reliable and suitable for process simulation and optimisation, the model for photosynthetic biomass growth has to be identified against the experimental data. It can be verified that mCRM shows identifiability issues if a parameter estimation is performed on the whole parameters vector $\widehat{\theta}$. The estimation of the model parameters is characterised by large confidence intervals for some parameters and more than one set of optimal values can be determined (i.e. the model is not uniquely identifiable). In order to overcome this problem both structural (or global) and practical identifiability of the model have been studied. First of all, the global identifiability has been verified using a differential algebra based method. Afterwards, the practical identifiability has been assessed through a sensitivity analysis and a model reparameterisation has been performed, as discussed in the following.

3.6.1 Global identifiability analysis

The first step for testing model identifiability is represented by global identifiability analysis. Global identifiability analysis is in fact a necessary condition to the practical identifiability and can provide the minimum number of observations required to identify an unique set of optimal parameter values. The two hypotheses, upon which structural identifiability analysis rely, are: i) complete absence of measurement errors, and ii) a perfectly accurate model structure. Those two assumptions refer to an ideal case and therefore, once a model is verified to be globally identifiable, practical identifiability, too, needs assessing. In the literature, several methods to

study the structural identifiability of non-linear models have been developed in the last two decades and are nicely reviewed in the work by (Miao et al., 2011).

A rather common approach is given by the series expansion method (e.g., Dochain et al. (1995) applied this approach to kinetic models of activated sludge respiration). It requires that functions representing the model are infinitely differentiable, as the method involve the calculation of arbitrary order derivatives. Several examples of application of this method can be found in literature. However, as pointed out by the authors, the series expansion method has a serious drawback: for high dimensional models high order derivatives are necessary and the resulting equations can easily become too complicated to solve. In fact, it was verified that in this case the series expansion method leads to an intractable problem. Thus, in order to overcome the difficulties related to high order derivatives calculation and the resolution of the resulting equations, a method based on differential algebra was taken into account. In particular, the software package DAISY (Bellu et al., 2007) has been used here. Being DAISY based on differential algebra techniques, it requires equations of the model to be written in the form of differential polynomials (Miao et al., 2011). In Equations 3.12 and 3.14 terms $(I - I_{cr})^{\alpha}$ and I^{α_2} have to be modified to be in form of differential polynomials. In particular is it possible to define two additional state variables, I_{α} and I_{α_2}, to be the Taylor series expansion of $(I - I_{cr})^{\alpha}$ and I^{α_2} respectively. Accordingly:

$$I_{\alpha} = (I - I_{cr})^{\overline{\alpha}} \left(\sum_{n=0}^{+\infty} \frac{(\alpha - \overline{\alpha})^n}{n!} \log^n (I - I_{cr}) \right) \tag{3.19}$$

$$I_{\alpha_2} = I^{\overline{\alpha_2}} \left(\sum_{n=0}^{+\infty} \frac{(\alpha_2 - \overline{\alpha_2})^n}{n!} \log^n (I) \right) \tag{3.20}$$

Where $\overline{\alpha}$ and $\overline{\alpha_2}$ are parameter values at which the series expansion has been considered. Being α and α_2 bounded between 0 and 1 in order to assure the global identifiability of the model several combinations of $\overline{\alpha}$ and $\overline{\alpha_2}$ have been considered. The idea was to analyse different part of the parametric space to have a good approximation of the original model in all the parameter space. In particular the following values have been considered for $\overline{\alpha}$ and $\overline{\alpha_2}$ [0.1; 0.3; 0.5; 0.7; 0.9]. The Taylor series were truncated at the second order for numerical reasons as an higher order Taylor expansion do not improve significantly the accuracy of the model approximation and lead to an higher computational cost. The system of equations 3.11 to 3.18, where terms 3.19 and 3.20 have been used to substitute $(I - I_{cr})^{\alpha}$ and I^{α_2} has been tested for global identifiability for each possible pair of values of $\overline{\alpha}$ and $\overline{\alpha_2}$.

Results indicate that a single experiment is not sufficient to identify a unique value of all the parameters. However, if (at least) three parallel experiments are carried out at different light intensities and both concentration and fluorescence parameters are measured during the experiments, the model is globally identifiable. This means that the experienced identifiability issues depend on practical identifiability. To tackle the problem, in the following a sensitivity

analysis and a reparamereterisation will be performed. Note that in several cases, where more complex models need considering, the analysis of the model practical identifiability may be the only test that can be carried out, since the verification of global identifiability may not be viable.

3.6.2 Sensitivity and correlation analysis

As the model is structurally identifiable, its practical identifiability will now be tested. The first step is represented by sensitivity analysis. The sensitivity of the i-th response to the k-th model parameter is defined as

$$q_{ik} = \frac{\partial y_i}{\partial \theta_k} \approx \frac{y_i^{'} - y_i}{\Delta \theta_k} \qquad i = 1, ..., N_M; \quad k = 1, ..., N_\theta \tag{3.21}$$

where y_i is the i-th measured responses predicted by the model, $y_i^{'}$ is the same response obtained from a perturbed value of the k-th parameter θ_k, and $\Delta\theta_k$ is the perturbation (N_M is the number of measured responses and N_θ is the number of model parameters). The principal goal of sensitivity analysis is to evaluate the impact of each parameter on the measured responses and to underline the presence of correlation among specific subsets of model parameters. In an over-parametrised model, near-zero sensitivity values would be obtained, leading to the non-identifiability of some subsets of model parameters. In the most desirable case, sensitivity profiles should be clearly distinct and far from being symmetrical (i.e. they should present a low mutual correlation). As the sensitivity analysis requires pre-specified parameters values, values by a preliminary parameter estimation have been considered. Perturbation $\Delta\theta_k$ was set equal to 1% of parameter values. First, the sensitivity profiles of biomass concentration will be taken into account. Although not shown here for the sake of conciseness, it was verified that all sensitivity profiles behave as exponential curves. As a consequence, the final value of sensitivities (value at 120 hours) is sufficient to analyse the system behaviour in different experimental conditions. The sensitivity profiles evaluated at four different light intensities are reported in Figure 3.6.

For light intensities under the critical value, the sensitivities related to parameters $k_{i,0}$ and k_r show opposite values. Above the critical value of light intensity, the sensitivity of k_r is the opposite of the sum of the sensitivities of $k_{i,0}$ and $k_{i,0}$. This suggests that it may be difficult to exploit biomass concentration measurements to identify $k_{i,0}$ and k_r.

Another critical aspect is related to the fact that parameters k_a and k_p exhibit a very similar sensitivity at low light conditions (availability of growth data at high light intensities may be necessary for a robust estimation of these parameters). Also it should be noticed that the parameter related to photoacclimation, k_c, and the parameter representing the maintenance factor in the dark, M, exhibit a very similar (and low) sensitivity in all light conditions.

Fluorescence profiles (q and q_L) are also considered. As the dynamics of PSUs are fast with

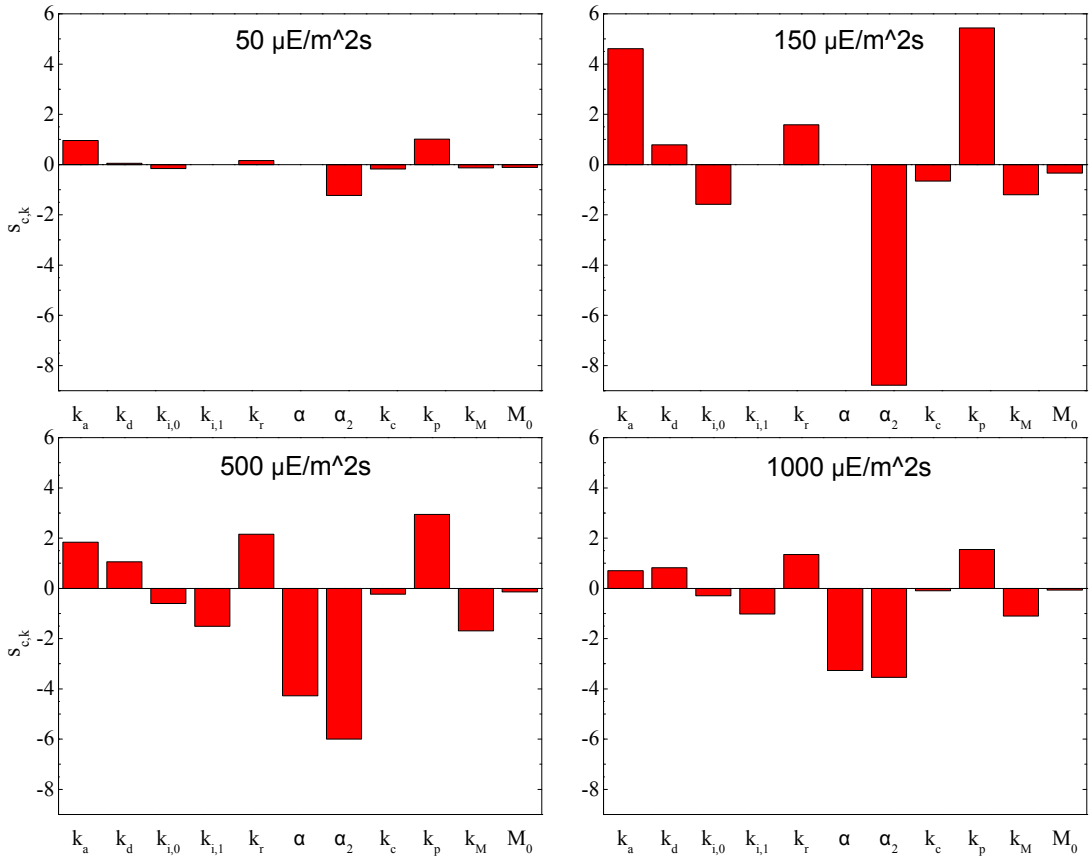


Figure 3.6: *Final values of dynamic sensitivities for mCRM, evaluated at different light intensities.*

respect to the sampling time, and our measurements are steady state measurements, only the steady state values of sensitivities will be reported. In Figure 3.7(a) and 3.7(b) the values of the sensitivities of q and q_L are reported for four different light conditions. The sensitivities of parameters α_2, k_p, k_c, k_M and M_0 are not reported, since they are zero for both fluorescence measurements (they are related to biomass growth and are not concerned with the oxidation state of the PSUs).

Considering the steady state sensitivity values of the fluorescence measurements, the following critical aspects can be noticed:

- parameters k_a and k_d do not affect the dark fluorescence measurements, while they are completely anticorrelated if light fluorescence measurements are available, showing opposite sensitivities in all light conditions;
- parameters $k_{i,0}$, $k_{i,1}$, k_r and α are not affected by light fluorescence measurements;
- parameter k_r is anticorrelated with $k_{i,0}$, under the critical value of light intensity, and with $k_{i,1}$, above the critical value of light intensity;
- the sensitivity of $k_{i,0}$ is always quite small. The results suggest that a model reparameterisation may be necessary to help tackling the issue (Meshkat et al., 2009).

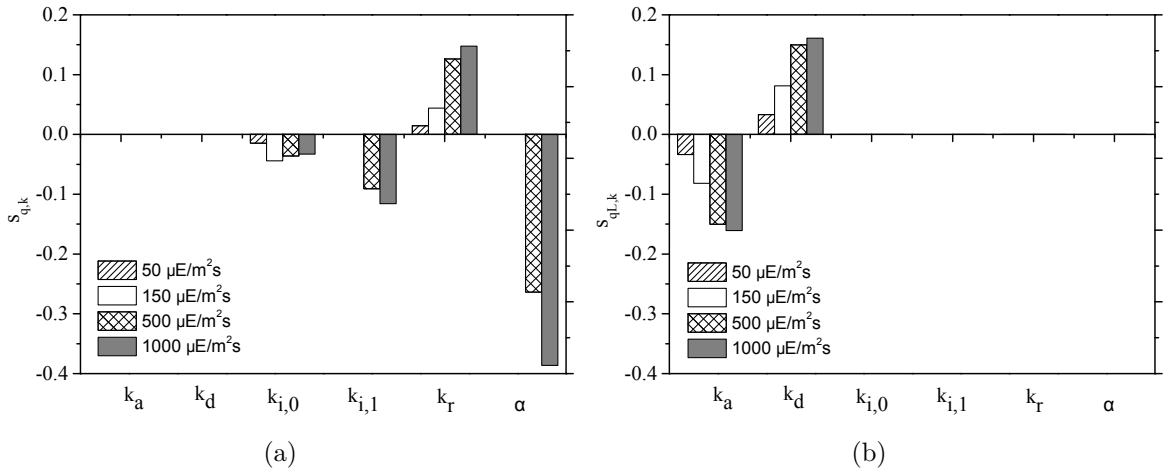


Figure 3.7: *Final values for dynamic sensitivities for the mCRM model evaluated at different light intensities considering literature values for model parameters.*

3.6.3 Reparameterisation of Camacho Rubio model

As the fluorescence data considered here are measurements of the PSU oxidation state under steady state conditions, an analytical expression for fluorescence measurements can be derived from Equation 3.11 to 3.13:

$$\frac{q}{q_{max}} = \frac{a_t - a_3}{a_t} = \begin{cases} 1 - R_0 I & \text{if } I \leq I_{cr} \\ \dfrac{1 - R_0 I_{cr}}{R_1 \left(I - I_{cr}\right)^{\alpha} + 1} & \text{otherwise} \end{cases} \tag{3.22}$$

$$q_L = \frac{a_1}{a_1 + a_2} = \frac{1}{R_2 I + 1} \tag{3.23}$$

with $R_0 = k_{i,0}/k_r$, $R_1 = k_{i,1}/k_r$ and $R_2 = k_a/k_d$. The interesting aspect is that only four parameters affect the steady state values of the fluorescence measurements: the three ratios R_0, R_1, R_2 and parameter α. This suggests that, in order to have a practically identifiable model, the values of two parameters affecting the PSU dynamics have to be fixed if no dynamics fluorescence data can be included in the data set. It was chosen to fix the values of k_r and k_d, as those parameters represent the rate constant of the recovery and de-excitation processes, respectively. An approximated estimation can be obtained, considering the timescales of the processes involved: according to literature (Han et al., 2000a), values of 100 s^{-1} for k_d and of $2.22 \cdot 10^{-4} s^{-1}$ for k_r were assumed. The very low sensitivity to available measurements and the high correlation of k_c and M_0 make it impossible to identify the two parameters. However, in the case of M_0, previous experiments suggest it to be between 5% and 10% of the maximum growth rate (Ryther, 1955). Here we set $M_0 = 1.5 \cdot 10^{-3} h^{-1}$ ($\approx$ 8% of the maximum growth rate). For parameter k_c the preliminary estimation suggests that a 'small' value is required for a good description of data. Since in Equation 3.14 we have that $I^{\alpha_2} \gg 1$, we verified that $k_c = 1$

is a good approximation for representing all experimental conditions. After reparametrisation, the vector of parameters to estimate is $\widehat{\theta}^* = [R_0, R_1, R_2, k_p, k_M, \alpha, \alpha_2]$.

3.7 Results and discussion

In the first part of this section the mCRM parameter estimation results will be presented and discussed. In the second part two additional growth curves and two new measurements of dark fluorescence will be used to validate the model.

3.7.1 Model calibration

In the estimation procedures, parameters are normalised with respect to the initial values obtained by the preliminary parameter estimation, to increase numerical robustness. The results of parameters estimation are reported in Table 3.6, along with the confidence intervals and the t-values (for a statistically precise estimation of a model parameter the t-value has to be greater than a reference t-value). The t-value statistic shows that the parameter values are estimated in a statistically satisfactory way.

Table 3.6: *Estimated values of parameters of the reparametrised mCRM, normalised values (with respect to the initial values) of the parameters, confidence intervals and t-student values for parameters. The reference t-value is equal to 1.67.*

Parameter	Estimated value	Normalized value	95% conf int	t-value 95%
R_0	4.93×10^{-4}	1.12	0.21	5.20
R_1	1.34×10^{-2}	0.65	0.21	3.70
R_2	3.30×10^{-3}	2.77	0.41	6.72
k_p	1.48×10^{-6}	0.41	0.12	3.35
k_M	1.12×10^{-1}	1.96	0.91	2.15
α	4.50×10^{-1}	0.45	0.048	9.52
α_2	3.00×10^{-1}	0.30	0.078	3.89

The profiles of fluorescence, predicted after the identification of the reparametrised mCRM, are reported in Figure 3.8. We can observe that the model correctly fits both the dark fluorescence profile (Figure 3.8(a)) and the light fluorescence measurements (Figure 3.8(b)). Biomass growth profiles are rather well represented by the model as illustrated in Figure 3.9, where the six different illuminating conditions are represented.

In Figure 3.10 the growth rate constant predicted by the model is reported along with the experimental value (i.e. the value obtained fitting the experimental data of growth during exponential phase with an exponential curve). We can observe that for all light intensities at which an experiment was carried out, the model well describes the experimental values of growth rate constant. This suggests that, thanks to the fundamental input of the fluorescence parameters in illuminated cells (q_L), the model is capable of reproducing with sufficient accuracy the basic processes of photosynthesis: photochemistry, light damage and also energy dissipation.

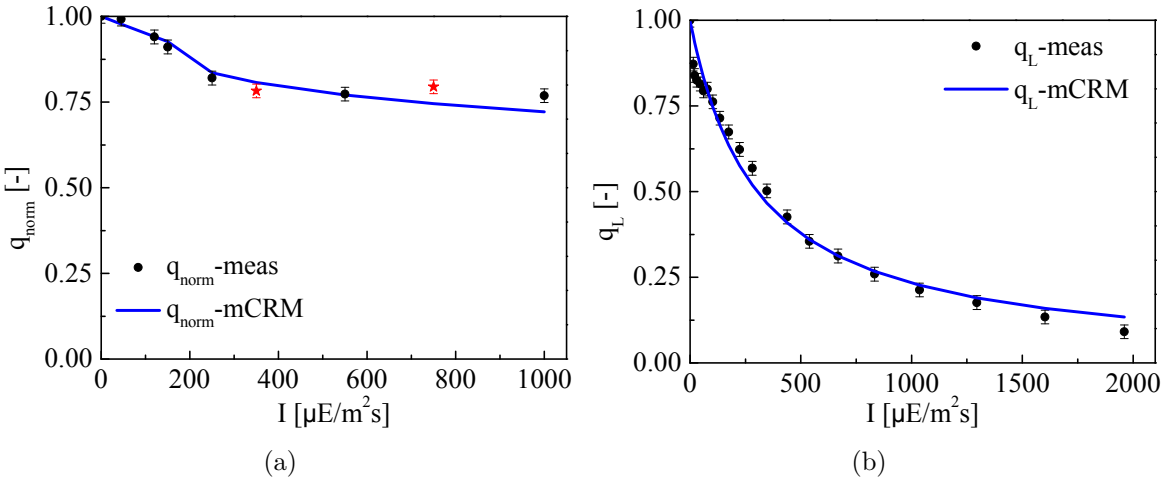


Figure 3.8: *Measurement (black circles) and predicted values of (a) q and (b) q_L. Blue solid lines represent the profiles predicted by the mCRM model. Red stars in panel (a) represent the experimental data used for model validation.*

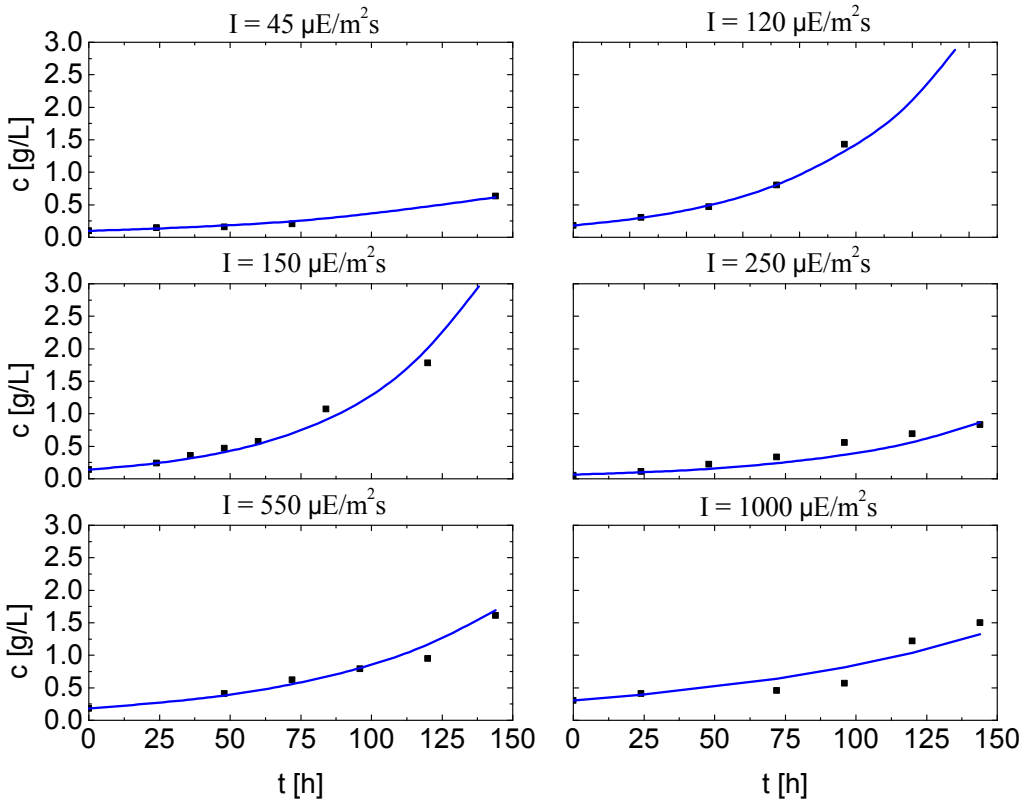


Figure 3.9: *Biomass concentration profiles at different light intensities predicted by the modified Camacho Rubio model. Black circles represent the experimental measurements*

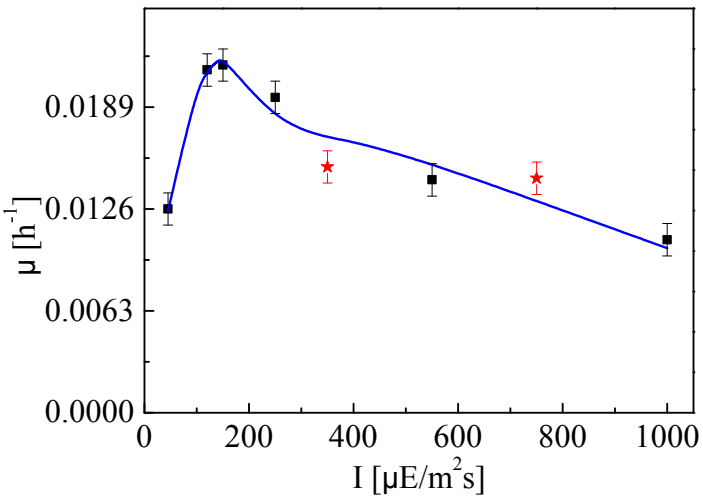


Figure 3.10: *Growth rate constant predicted by the modified Camacho Rubio model (solid line) and experimental values of the growth rate constant (black circles). Red stars represent the experimental values of the growth rate constant for the experiments used in the model validation*

From a statistical point of view, the predicted profiles have a χ^2 value of 117.7 (whereas $\chi^2 = 890.1$ in the case of EPM, and $\chi^2 = 301.6$ for CRM).

3.7.2 Model validation

In order to validate the model, two experiments, not used for model calibration, will be taken into account. In particular two growth curves at 350 and 750 μE/m^{2}s have been considered. For both cultures also the value of q has been measured and was exploited for model validation. In Figures 3.8(a) and 3.10 the validation points are represented by red stars. In Figure 3.11 the growth curves have been reported.

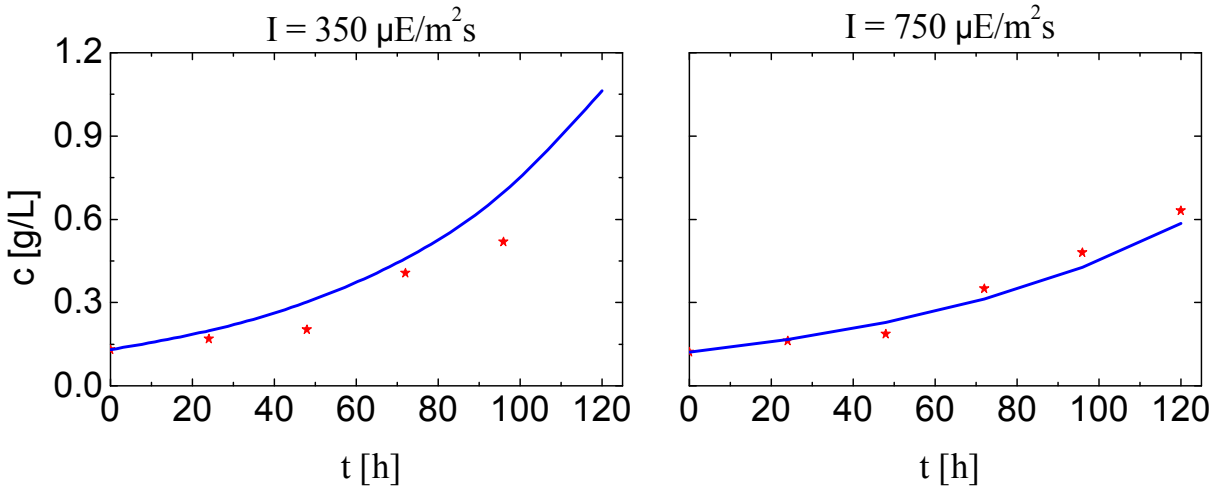


Figure 3.11: *Biomass concentration profiles at different light intensities predicted by the modified Camacho Rubio model. Red stars represent the experimental measurements used for the model validation*

We can observe that the predictions are accurate for the parameter q. Also the growth

curves are predicted in a sufficiently accurate way, although the growth at $350\mu E/m^2s$ is slightly overestimated by the model, while the growth at $750\mu E/m^2s$ is somewhat underestimated in the final part. From a statistical point of view, validation leads to χ^2 value of 226.23.

An informative use of the model involves estimating the energy needed for cell maintenance at different light intensities, predicting an increase in maintenance energy requirements at higher irradiance, likely caused by a larger repair rate in damaged photosystems. The latter becomes clear in plotting the ratio between the photosynthesis and maintenance rates (Figure 3.12). We can observe from Figure 3.12 that the optimal growth intensity, i.e. the light intensity at which the growth is maximum, correspond to a minimum in the ratio between maintenance term and gross growth rate (sum of the net growth rate μ with maintenance). Even though the model may not be generally applicable to any microalgae culture since it fails to account for key mechanisms such as nutrients limitation, it nevertheless provides a means for quantifying the weight carried by photoinhibition, clearly suggesting that this mechanism is likely to have a major impact on microalgae productivity at high light irradiance.

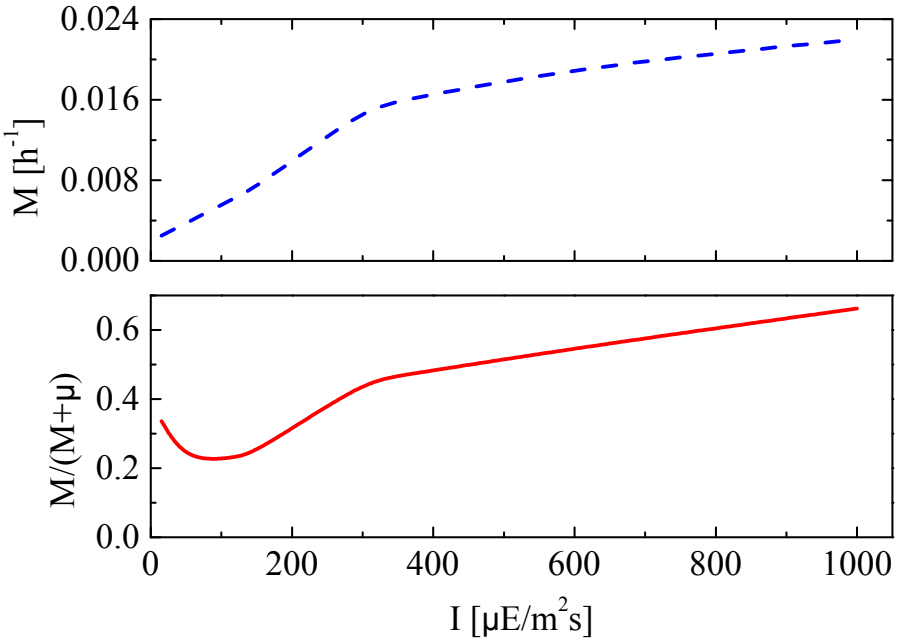


Figure 3.12: *The upper plot shows the dependence of the maintenance factor (expressed in* h^{-1}*) from light intensity. The lower plot represent the ratio between the maintenance and the gross growth rate (sum of net growth rate with maintenance) dependence from light intensity.*

3.8 Conclusion

A literature model (Rubio et al., 2003) was selected and modified to describe microalgae growth through a rigorous identification procedure. The estimation of the model parameters was performed considering experimental data (growth profiles and fluorescence measurements) on Nannochloropsis Salina. Results show that the developed model well represents biomass growth

over a wide range of light intensities. The modified model was also able to reproduce fluorescence measurements, including the light profile of PSU saturation. This suggests that the proposed model is accurate enough to represent all major processes of photosynthesis, photochemistry, PSU damage and energy dissipation. While results in reproducing experimental data are fully satisfactory, it should be underlined that algae growing in an industrial scale photobioreactor are exposed to different conditions. In particular light is not homogenously distributed because of cells shading, and illumination intensity is not constant because of diurnal changes and cells mixing. Finally nutrient availability can also be limiting.

In the next chapter a semi-mechanistic model of chlorophyll fluorescence will be presented and discussed. This model will allow us to critically evaluate the relationship between q and inhibited PSUs. Although it is used in the literature as a proxy of inhibited PSUs in similar studies (e.g.: Wu and Merchuk (2001)) the linear relationship is indeed still matter of discussion. The analysis of fluorescence model will show how this hypothesis can be linked to the inhibition-related quenching. Other two limitations of the model are: (i) the acclimation effect of reducing the number of PSUs is actually only one of several effect induced from light adaptation and (ii) the energy dissipation NPQ is only described in a static way while its activation is light and time dependent. However, the simple way we choose to represent those effects are related to the type of data we aimed to describe and are enough to represent the overall light effects on growth. In the following chapter we will show how it is possible to better represent those phenomena exploiting the fluorescence measurements.

Chapter 4

A model of chlorophyll fluorescence in microalgae integrating photoproduction, photoinhibition and photoregulation

This chapter[1] presents a mathematical model capable of delivering a quantitative prediction of the state of the photosynthetic apparatus of microalgae in terms of their open, closed and damaged reaction centres under variable light conditions. This model combines the processes of photoproduction and photoinhibition in the Han model with a novel mathematical representation of photoprotective mechanisms, including qE-quenching and qI-quenching. For calibration and validation purposes, the model can be used to simulate fluorescence fluxes, such as those measured in PAM fluorometry, as well as classical fluorescence indexes. A calibration is carried out for the microalga *Nannochloropsis gaditana*, whereby 9 out of the 13 model parameters are estimated with good statistical significance using the realized, minimal and maximal fluorescence fluxes measured from a typical PAM protocol. The model is further validated by considering a more challenging PAM protocol alternating periods of intense light and dark, showing a good ability to provide quantitative predictions of the fluorescence fluxes even though it was calibrated for a different and somewhat simpler PAM protocol. Applications of the model include the ability to predict PI-response curves solely based on PAM fluorometry togheter with the long term goal of combining it within first-principles models describing the flow and light attenuation for design and operation of large-scale microalgae production systems.

[1]Part of the work of this chapter has been published in the work by Nikolaou et al. (2014a) and Nikolaou et al. (2015)

4.1 Motivation

Mathematical models are commonly used in chemical engineering for process optimisation (Biegler et al., 2014), but their application to microalgae culture is made particularly arduous by the large number of governing phenomena, acting on multiple time scales ranging from milliseconds to days. *Photoproduction* refers to all the processes from photons utilization to CO_2 fixation and occurs withing milliseconds (Williams and Laurens, 2010). *Photoinhibition* is the observed loss of photosynthetic production due to excess or prolonged exposure to light and acts on time scales of minutes to hours (Long et al., 1994). *Photoregulation*, often referred as NPQ is the ability of photosynthetic organisms to dissipate excess excitation energy as heat, thereby protecting the photosynthetically active components, and occurs within minutes (Müller et al., 2001). *Photoacclimation* is the ability of microalgae to adjust their pigment content and composition under varying light and nutrient conditions and acts on time scales of hours to days (MacIntyre et al., 2002).Finally, the mechanisms involved in nutrient internalization and their metabolism into useful products occur within hours to days as well (Falkowski and Raven, 1997).

Chlorophyll fluorescence is a powerful tool for the analysis of the aforementioned processes, which has led to important discoveries over the past 40 years. Today's state-of-the-art equipment, such as PAM, can implement complex protocols with great measurement accuracy. In order to fully exploit this capability, experimental protocols have been developed that relate the measured fluorescence fluxes with key photosynthetic parameters such as the quantum yield of photosynthesis, the photosynthetic apparatus activity, and the NPQ activity. Even if the level of understanding of the various fluorescence parameters has significantly increased the past years, little effort has been devoted to the development of dynamic models that associate the operation of the photosynthetic machinery with fluorescence measures. In most of the literature the fluorescence indexes are only used to compare different cultures in a qualitative way, but only few contributions try to exploit the quantitative information that can be extracted from those kind of measurements (Wu and Merchuk, 2001).

The main objective of this chapter is to develop a mathematical model that can predict the fluorescence fluxes in terms of the photosynthetic mechanisms occurring inside the chloroplasts. This model builds upon the widely-accepted state-transition model proposed by Han (2002) for predicting photoproduction and photoinhibition. An extension of this model in the form of a semi-empirical expression is proposed in order to encompass a particular type of photoregulation, namely qE quenching. Moreover, the chlorophyll-a fluorescence flux is expressed based on the work by Huot and Babin (2010), for a specific type of PSII arrangement, the so-called lake model (Kramer et al., 2004). The novelty and originality of the model lies in the way the states of the PSUs, as given by the (extended) Han model, are related to the measured fluorescence parameters, and how it does so by accounting for qE quenching.

The remainder of the chapter is organized as follows. The experimental measurements used

to develop, calibrate and validate the model are described in Section 4.2. In Section 4.3 the fluorescence model is presented and discussed. Next, the calibration and validation results are reported, showing an excellent agreement between experimental data and model predictions. Some final remarks will conclude the chapter.

4.2 Material and methods

The microalga *Nannochloropsis gaditana* (CCAP, strain 849/5) was grown in a sterile, filtered F/2 medium, using sea salts (32 g/L) from Sigma, 40 mMTris HCl, pH 8 and Sigma Guillard's (F/2) marine water enrichment solution. Growth experiments were conducted in the multi-cultivator MC 1000-OD system (Photon Systems Instruments, Czech Republic), with daily measurements of the growth rate via changes in optical density OD 720 using spectrophotometry. The suspension culture was continuously mixed and aerated by bubbling air, maintained at a temperature of 21 °C, and subject to a constant light intensity of 100 μE/m^{2}ssupplied by an array of white LEDs. Samples were harvested from the multicultivator after 5 days (late exponential phase), so that the microalgae are acclimated to these conditions, yet still actively growing and not experiencing nutrients depletion. A pre-culture was also grown at 100 μE/m^{2}sin glass bottles of 0.25 L under a continuous airflow, enriched with 5% CO_2. At the exponential phase, the pre-culture was centrifuged and re-suspended in fresh medium to reach a final concentration of 9×10^6 cells/ml, before its introduction in the multi-cultivator. All the fluorescence measurements were performed using a Dual PAM (Walz, Germany), after a dark adaptation period of 20 minutes, by exposing the microalgae samples to variable actinic light intensities in time intervals of 60 seconds. Before switching-on of the actinic light and during the final 2 s of each interval, a saturating light pulse at 6000 μE/m^{2}swas applied during 0.6 s, followed by a dark period (actinic light off) of 1.4 s; measurements were recorded before and after the saturating pulses and after the dark periods, which correspond to $F^{'}$, $F_m^{'}$ and $F_0^{'}$ respectively.

Two separate experiments were performed for the purpose of model calibration 4.4.1 and validation 4.4.2. Both light protocols are reported, with the corresponding fluorescence flux data, in Table 4.1. The simulations of the fluorescence model were conducted in the modelling environment gPROMS (Process System Enterprise, 2012). The calibration too was performed in gPROMS using maximum likelihood estimation and statistical confidence analysis (Walter and Pronzato, 1997), in order for the model predictions to match the measured fluorescence fluxes $F^{'}$, $F_m^{'}$ and $F_0^{'}$. Due to lack of further information regarding the precision and accuracy of the PAM fluorometer, a 1% standard deviation was assumed for the measured fluorescence fluxes to determine the parameters and estimate their confidence intervals.

Table 4.1: *Actinic light and fluorescence flux measurements for the calibration experiment.*

	Calibration	Validation	Calibration			Validation		
Time [s]	I μE/m^{2}s		$F^{'}$ [V]	$F^{'}_m$ [V]	$F^{'}_0$ [V]	$F^{'}$ [V]	$F^{'}_m$ [V]	$F^{'}_0$ [V]
0	0	0	0.077	0.223	0.078	0.09	0.221	0.09
60	14	14	0.083	0.224	0.078	0.094	0.222	0.09
120	21	21	0.082	0.227	0.078	0.094	0.225	0.091
180	45	1602	0.083	0.226	0.078	0.064	0.067	0.047
240	78	1960	0.084	0.222	0.078	0.053	0.055	0.04
300	134	45	0.087	0.214	0.077	0.087	0.161	0.078
360	174	78	0.088	0.209	0.076	0.09	0.174	0.081
420	224	1036	0.088	0.198	0.075	0.051	0.057	0.041
480	281	1295	0.087	0.18	0.072	0.046	0.049	0.037
540	347	134	0.083	0.16	0.069	0.079	0.136	0.072
600	438	174	0.079	0.137	0.064	0.08	0.135	0.071
660	539	1602	0.073	0.116	0.059	0.048	0.05	0.038
720	668	1960	0.067	0.098	0.054	0.042	0.043	0.033
780	833	45	0.061	0.082	0.049	0.075	0.127	0.069
840	1036	78	0.057	0.07	0.044	0.082	0.142	0.073
900	1602	134	0.053	0.061	0.041	0.08	0.137	0.072
960	1602	1960	0.05	0.055	0.038	0.047	0.048	0.036
1020	1960	1960	0.047	0.051	0.036	0.042	0.042	0.033
1080	14	0	0.07	0.146	0.066	0.063	0.102	0.061
1140	14	0	0.07	0.154	0.067	0.064	0.112	0.064
1200	14	0	0.071	0.16	0.069	0.067	0.12	0.067

4.3 A dynamic model of fluorescence in microalgae

This section presents a dynamic model of chlorophyll fluorescence that accounts for key photosynthetic processes having time scales up to an hour. Specifically, the model encompasses the processes of photoproduction, photoinhibition and photoregulation, but neglects the changes in photoacclimation state. The photoacclimation effects will be assessed in Chapter 5.

In PAM fluorometry, the fluorescence flux F [V] emitted by a microalgae sample of volume V [m^3] and chlorophyll concentration chl [$g_{chl}\,m^{-3}$] can be modelled as (Huot and Babin, 2010):

$$F = I_m\,\sigma\,chl\,\Phi_f\,(1-Q)\,V\,\lambda_{\text{PAM}}\,, \tag{4.1}$$

where Φ_f stands for the quantum yield of fluorescence [$\mu E\,\mu E^{-1}$]; σ, the total cross section [$m^2\,g_{chl}^{-1}$]; I_m, the measuring light intensity [μE/m^{2}s]; Q is a dimensionless parameter describing the percentage of fluorescence absorbed by the sample; and λ_{PAM} [$V\,s\,\mu E^{-1}$] is a gain parameter aligning the voltage output of a PAM fluorometer with the actual fluorescence flux. In particular, as chl and Q remain constant for a given photoacclimation state, it is convenient to lump all the constant terms into a single parameter $S_F := I_m\,chl\,(1-Q)\,V\,\lambda_{\text{PAM}}$, giving

$$F = S_F\,\sigma\,\Phi_f\,. \tag{4.2}$$

In the sequel, we use the Han model to represent the effects of photoproduction and photoinhibition on the fluorescence flux (Section 4.3.1). Next we formulate two alternative

Han model extensions in order to encompass photoregulation effects and we discriminate the two alternatives selecting the best one (Section 4.3.2). Then, we analyze the properties of the resulting model (Section 4.3.3). Finally we calibrate and validate the final model.

4.3.1 Han model

The model developed by Han (2002) and originating in the works of Kok (1956) and Eilers and Peeters (1988) is based on the concept of photosynthetic unit (PSU), first introduced by Gaffron and Wohl (1936) to represent the physical entity responsible for the production of one O_2 molecule. In this conceptual representation, each PSU is comprised of one RCII and its associated LHC, and the chloroplasts are regarded as PSU arrays. Equivalent state models have been extensively used to mathematically represent biomass productivity and can prove particularly useful in studies of fluctuating light effects (Rubio et al., 2003). With respect to the models considered in the Chapter 3 the Han model makes use of a parametrisation that gives at all the parameters a precise biological meaning as we will show in the following. In particular, the Han model parametrisation will be particularly effective to link the fluorescence expression with the state model.

The description of photoproduction and photoinhibition in the Han model assumes that the RCII of a PSU can be in either one of three states, namely open (A), closed (B) or damaged (C). An RCII in state A is ready to accept an electron; in state B, it is already occupied by electrons; and in state C, it is non-functional. As depicted in Figure 4.1, each RCII can transit from one state to another depending on the light irradiance I, with processes described by first-order kinetics. Photoproduction is described by the transition from A to B, while the reverse transition from B to A represents relaxation of the RCII; photoinhibition, on the other hand, corresponds to the transition from B to C, while the reverse transition from C to B describes repair of the damaged RCII by enzymatic processes.
The equations in the Han model describe the dynamics of the fractions of open, closed and damaged RCIIs in the chloroplasts, denoted by $A(t)$, $B(t)$ and $C(t)$, respectively:

$$\begin{aligned} \dot{A} &= -I\,\sigma_{\mathrm{PS2}}\,A + \frac{B}{\tau} \\ \dot{B} &= I\,\sigma_{\mathrm{PS2}}\,A - \frac{B}{\tau} + k_r\,C - k_d\,\sigma_{\mathrm{PS2}}\,I\,B \\ \dot{C} &= -k_r\,C + k_d\,\sigma_{\mathrm{PS2}}\,I\,B\,. \end{aligned} \tag{4.3}$$

Here, σ_{PS2} denotes the effective cross section [$\mathrm{m}^2\,\mu\mathrm{E}^{-1}$]; τ, the turnover time [s]; k_d, the damage rate constant [-]; and k_r, the repair rate constant [s^{-1}]. Moreover, $A(t) + B(t) + C(t) = 1$ at all times.

Several expressions of the fluorescence quantum yield Φ_f in (4.2) as a function of the PSU states A, B and C have been proposed depending on the LHC-RCII configuration. They

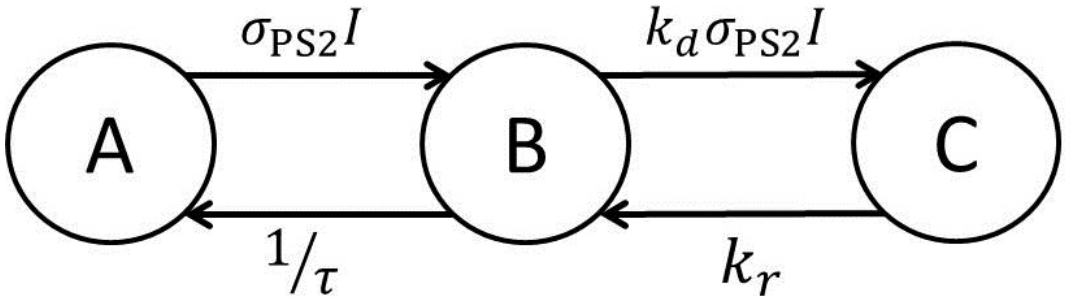


Figure 4.1: *Schematic representation of the Han model*

typically involve the parameters Φ_f^A, Φ_f^B and Φ_f^C representing the fluorescence quantum yields of an RCII in state A, B or C, respectively (Huot and Babin, 2010). The one configuration used subsequently is the so called lake model (Kramer et al., 2004) expressing Φ_f as the harmonic mean of Φ_f^A, Φ_f^B and Φ_f^C[2].

$$\Phi_f := \frac{1}{\frac{A}{\Phi_f^A} + \frac{B}{\Phi_f^B} + \frac{C}{\Phi_f^C}} \,. \tag{4.4}$$

In analogy to parallel electrical circuits, this configuration assumes that all RCIIs are connected to a common LHC and thus compete for the incoming excitation energy. Naturally, other types of LHC-RCII configurations can be used in the fluorescence model if desired.

Besides Φ_f, the total cross-section σ in (4.2) can be related to the parameter σ_{PS2} in the Han model. In a first step, σ is related to the so-called optical cross section, σ_{PSU} $[m^2\,\mu mol(O_2)^{-1}]$, as

$$\sigma = \sigma_{PSU}\, N \,, \tag{4.5}$$

with N the number of PSUs $[\mu mol(O_2)\, g_{chl}^{-1}]$, which remains constant for a given photoacclimation state. In a second step, σ_{PS2} can be related to σ_{PSU} as (Falkowski and Raven, 1997):

$$\sigma_{PS2} = \nu \Phi_p^A\, \sigma_{PSU} \,, \tag{4.6}$$

where Φ_p^A [-] denotes the quantum yield of photosynthesis of an open RCII, which is equal to the realized quantum yield of photosynthesis Φ_{PS2} in the case that $A = 1$ (see Section 4.3.3 below for an expression of Φ_p^A); and ν is a conversion factor that aligns the electrons required to dissociate one molecule of water. If we consider the water dissociation reaction ($2H_2O + 4e^- \rightarrow O_2 + 4H^+$) a minimum theoretical value of $\nu = 4\ \mu E\,\mu mol(O_2)^{-1}$ can be derived.

[2] a discussion about the Φ_f expression is presented in Appendix A

4.3.2 Accounting for photoregulation

An important limitation of the Han model in the context of PAM fluorometry is that some of its parameters may vary on the time scale of minutes due to certain NPQ regulation mechanisms being activated. Two types of NPQ mechanisms are accounted for in this work, namely qI-quenching and qE-quenching, which are often seen as the major contributors to fluorescence quenching on the time scales of interest (Horton and Hague, 1988). qE-quenching is activated at high light irradiance by low thylakoid lumenal pH (Bilger and Björkman, 1990); it evolves within minutes and can result in up to 90% reduction in fluorescence (Huot and Babin, 2010). qI-quenching is linked to photoinhibition, according to the biological hypothesis that damaged RCII can trap and dissipate excited electrons as heat; it typically evolves in a time scale of minutes to hours and can be responsible for up to 40% reduction in fluorescence (Falkowski et al., 1994).

We start by noting that only qE-quenching requires further consideration as qI-quenching is already accounted for through the dependence of Φ_f on the fraction C of damaged RCIIs in (4.4). Since qE-quenching in the dark is negligible and varies with the light irradiance via the change in lumenal pH, we introduce a conceptual qE activity reference function α_{SS} taking values in the range $[0, 1]$ and increasing with I, from the level $\alpha_{SS} = 0$ at $I = 0$. After consideration of experimental measurements of the NPQ index (2.14) as a function of I (Kramer et al., 2004), we choose to formulate α_{SS} as a sigmoid (Hill) function of I:

$$\alpha_{SS}(I) := \frac{I^n}{I_{\mathrm{qE}}{}^n + I^n}\,, \tag{4.7}$$

where I_{qE} [$\mu\mathrm{E}^{-1}$] represents the irradiance level at which half of the maximal qE activity is realized ($\alpha_{SS} = 0.5$); and n [–] describes the sharpness of the transition, approaching switch-like behavior as n becomes larger. In addition, we describe tracking of the qE activity reference α_{SS} by the actual qE activity level $\alpha(t)$ as a first-order process:

$$\dot{\alpha} = \xi\,(\alpha_{SS}(I) - \alpha)\,, \tag{4.8}$$

where ξ [s^{-1}] denotes the rate of NPQ adaptation, which shall be assumed constant here on the time scales of interest.

The activity variable formulated above needs to be included in the definition of the fluorescence flux (Equation 4.2 in order to account for the qE-quenching. Since S_F is constant for a given photoacclimation state, the effect can come through a dependence either on σ or on Φ_f. We decided to assume both alternatives thereby formulating two distinct models; in the first, qE-quenching is affecting the absorption efficiency of the LHC through variability in σ; in the second qE-quenching comes in effect through variability in the fluorescence quantum yields of A, B and C.

Variability in σ

The effect of qE-quenching comes forward through the variation in photon absorption by the LHC. A natural way of modelling this effect is therefore in terms of the total cross section, σ, the photosynthetic parameter capturing the light absorption effectiveness most directly. More specifically, we assume that σ is bounded between a maximum value, σ_{max}, and a minimal value, σ_{min}, which are observed when the NPQ energy dissipation is, respectively, the lowest and the highest:

$$\sigma = \sigma_{max}\,(1-\alpha) + \sigma_{min}\alpha \tag{4.9}$$

Combining 4.9, 4.5 and 4.6 yields the following expression of σ_{PS2} in terms of the photoregulation parameters:

$$\sigma_{\mathrm{PS2}} = \frac{\Phi_{\mathrm{p}}^{A}}{N}\,(\sigma_{max}\,(1-\alpha) + \sigma_{min}\alpha) \tag{4.10}$$

which can be used in the Han model 4.3 equations.

Variability in Φ_f

An alternative formulation of qE quenching comes forward by considering analytical expression for Φ_f^A, Φ_f^B and Φ_f^C. Following Huot and Babin (2010); Kitajima and Butler (1975); Oxborough and Baker (1997), we express the fluorescence quantum yields as:

$$\Phi_f^A = \frac{1}{1+\eta_P+\eta_D+\eta_{qE}}\,, \qquad \Phi_f^B = \frac{1}{1+\eta_D+\eta_{qE}}\,, \qquad \Phi_f^C = \frac{1}{1+\eta_I+\eta_D+\eta_{qE}}\,, \tag{4.11}$$

where the parameters η_P, η_D, η_{qE} and η_I represent, respectively, the rates of photoproduction, basal thermal decay in dark-adapted state, qE-quenching and qI-quenching, all relative to the rate of fluorescence; that is, these four parameters are dimensionless. Observe that Φ_f^B does not depend on η_P as a closed RCII cannot support photoproduction, and Φ_f^C depends on η_I instead of η_P in order to account for the effect of qI-quenching. Moreover, we capture the effect of qE-quenching by varying η_{qE} with respect to the qE activity level α as:

$$\eta_{qE} = \overline{\eta}_{\mathrm{qE}}\,\alpha\,, \tag{4.12}$$

with $\overline{\eta}_{\mathrm{qE}}$ a parameter describing the maximum rate of energy dissipation. Note that a linear relationship is assumed as a first approximation following the considerations by Oxborough and Baker (2000), but the model is readily extended to other, more complex dependencies. An important difference between the two ways of representing the qE-quenching effect is that the variability on σ affects the fluorescence independently from the RC oxidation state while the variability on Φ_f has a stronger effect on closed RC according to 4.11.

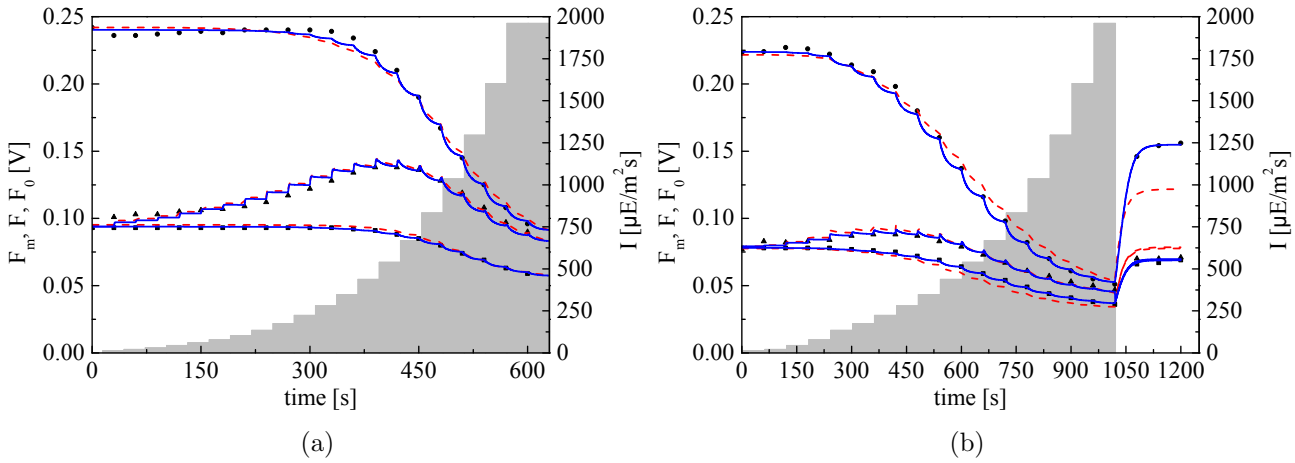


Figure 4.2: *Circles, triangles and squares represent the measured values of F_m', F_0' and F' respectively. The continuous line refers to the variable Φ_f model, while the dashed line refers to the variable σ model. Grey shaded area represent the light profiles.*

Model discrimination

Having two alternative modelling hypothesis the first objective is to discriminate between them in order to choose the most suitable one to represent our system. Two different experiments have been considered for model discrimination. The first experiment considers a PAM protocol where the light follows an increasing light intensity step profile with steps of 30 seconds constant light. In the second experiment each light step is 60 seconds long and at the end three minutes of dark have been added to evaluate the relaxation of the qE quenching. The results are reported in Figure 4.2: the continuous line represent the variable Φ_f model, the dashed line represent the variable σ model.

We can observe the experiment without a recovery phase can not be used to discriminate between the alternative models (Figure 4.2(a)). In fact, both model correctly represent the experimental data. On the contrary, the second experiment clearly indicates that the variable Φ_f model outperforms the variable σ model. We can observe from Figure 4.2(b) that the variable σ model is not able to represent the fast recovery of F_m' while the variable Φ_f model perfectly represent the experimental data. For this reason the variable σ model has been rejected and will not be discussed further.

4.3.3 Properties of fluorescence model

After model discrimination an expression of the fluorescence flux F can be obtained by substituting (4.11) and (4.12) back into (4.2):

$$F = \frac{S_F \sigma}{1 + \eta_D + \overline{\eta}_{\mathrm{qE}}\,\alpha + A\,\eta_P + C\,\eta_I}\,. \tag{4.13}$$

In PAM fluorometry, the fluorescence flux F in (4.13) corresponds to the light-adapted realized fluorescence flux $F^{'}$ (see Section 2.2.2). The remaining characteristic fluorescence fluxes F_0, F_m, $F_0^{'}$ and $F_m^{'}$ are obtained by specializing (4.13) with $A = 0$ and $B = 0$ for the maximal and minimal fluorescence fluxes, respectively, and with $\alpha = 0$ for the dark-adapted fluorescence fluxes. These expressions are reported in the left part of Table 4.2.

Table 4.2: *Expressions of PAM fluorescence fluxes (left part) and fluorescence indexes (right part).*

Flux	Expression
$F^{'}$	$\dfrac{S_F\sigma}{1+\eta_D+\overline{\eta}_{\mathrm{qE}}\alpha+A\eta_P+C\eta_I}$
$F_m^{'}$	$\dfrac{S_F\sigma}{1+\eta_D+\overline{\eta}_{\mathrm{qE}}\alpha+C\eta_I}$
$F_0^{'}$	$\dfrac{S_F\sigma}{1+\eta_D+\overline{\eta}_{\mathrm{qE}}\alpha+(1-C)\eta_P+C\eta_I}$
F_m	$\dfrac{S_F\sigma}{1+\eta_D+C\eta_I}$
F_0	$\dfrac{S_F\sigma}{1+\eta_D+(1-C)\eta_P+C\eta_I}$

Index	$\eta_P \neq \eta_I$	$\eta_P = \eta_I$
Φ_{PS2}	$\dfrac{A\eta_P}{1+\eta_D+\overline{\eta}_{\mathrm{qE}}\alpha+A\eta_P+C\eta_I}$	$\dfrac{A\eta_P}{1+\eta_D+\overline{\eta}_{\mathrm{qE}}\alpha+(1-B)\eta_P}$
q	$\dfrac{(1-C)\eta_P}{1+\eta_D+(1-C)\eta_P+C\eta_I}$	$\dfrac{(1-C)\eta_P}{1+\eta_D+\eta_P}$
q_L	$\dfrac{A}{A+B}$	$\dfrac{A}{A+B}$
q_{NPQ}	$\dfrac{\eta_{qE}+(C-C_0)\eta_I}{1+C_0\eta_I}$	

Mathematical expressions of the fluorescence indexes discussed in Section 2.2.3 follow readily from substitution of the foregoing PAM flux expressions. Two sets of expressions are reported in the right part of Table 4.2, corresponding to whether or not the assumption $\Phi_f^A = \Phi_f^C$ is made - or, equivalently, $\eta_I = \eta_P$. This assumption originates in the work of Maxwell and Johnson (2000), who argued that quenching related to damage in RCII does not cause a variation in the level of $F_0^{'}$ which imply that the fraction of incoming photons leading to photoproduction in an open RCII should be the same as the fraction of incoming photons dissipated as heat in a damaged RCII.

A number of comments are in order regarding the fluorescence index expressions:

- The realized quantum yield of photosynthesis, Φ_{PS2}, turns out to be a nonlinear function of the open, closed and damaged RCII fractions, irrespective of the assumption $\Phi_f^A = \Phi_f^C$. This result suggests that the usual hypothesis of a linear relationship between Φ_{PS2} and the fraction A of open RCIIs could be inaccurate, especially when the fraction B of closed RCIIs is small. An expression of the quantum yield of photosynthesis of an open RCII, defined as Φ_{p}^A earlier in (4.6), can also be derived from the expression of Φ_{PS2} in the special case that $A = 1$:
$$\Phi_{\mathrm{p}}^A = \frac{\eta_P}{1+\eta_D+\overline{\eta}_{\mathrm{qE}}\alpha+\eta_P}\,. \tag{4.14}$$

- The maximum quantum yield of photosynthesis, q, is a nonlinear function of the fraction C of damaged RCIIs in the dark-adapted sample in general, but this dependency becomes linear under the assumption that $\Phi_f^A = \Phi_f^C$. In other words, only if the quenching effect of

inhibited reaction centres perfectly compensate the loss of photoproduction capability, due to a smaller fraction of active reaction centres, the linear relationship assumed between q and inhibited PSU used in Chapter 3 and in literature work, such as Wu and Merchuk (2001), is valid.

- The photochemical quenching index, q_L, is found to be equal to the ratio of open-to-active RCIIs, which is in agreement with the considerations in Kramer et al. (2004).

Finally, we note that an expression of the fraction C of damaged RCIIs can be obtained as a function of the fluorescence indexes Φ_{L} and q_L in the form (see Section 2.2.3 for fluorescence indexes definition):

$$C = 1 - \frac{\Phi_{\mathrm{L}}(1 + \eta_D + \overline{\eta}_{\mathrm{qE}}\alpha - \eta_I)}{\Phi_{\mathrm{L}}(\eta_I - \eta_P) + q_L\eta_P}\,. \tag{4.15}$$

This relation is particularly useful from a practical standpoint as it allows predicting the level of damage of the photosynthetic apparatus based on experimental measurements of Φ_{L} and q_L, in combination with the qE activity level α predicted by (4.7)-(4.8). In particular, the latter equations are independent of the states of the PSUs due to the cascade structure of the fluorescence model. Similar expressions can be obtained for the fractions A and B of open/closed RCIIs by noting that $A = (1 - C)q_L$ and $B = (1 - C)(1 - q_L)$.

4.4 Results and discussion

This section presents the calibration results of the chlorophyll fluorescence model developed in Section 4.3 together with a validation analysis.

4.4.1 Model calibration

The chlorophyll fluorescence model developed in 4.3 comprises a total of 13 parameters, many of which have unknown values and thus need to be estimated. The light protocol and fluorescence flux measurements used for purpose of model calibration are shown on 4.3 (grey-shaded area and points with error bars, respectively). The first part of the experiment shows a gradual increase of the actinic light intensity from 0 to 1960 $\mu\mathrm{E/m^2sin}$ stages of 60 s, before the switching-off of the actinic light around 1000 s until the final time of 1200 s. The corresponding data are reported in Table 4.1.

Not all 13 parameter values can be estimated with high confidence from this data set, as certain parameters are insensitive or turn out to be highly correlated, if at all identifiable. Model reduction techniques could be used in order to arrive at a simpler model, but this would entail loss of physical meaning for (part of) the states and/or parameters and so was not considered here. After solving multiple instances of the parameter estimation problem for various subsets of parameters, it was found that nine parameters can be confidently estimated by keeping the following four parameters τ, k_r, N and η_D constant:

- The parameter τ representing relaxation of the closed RCIIs in the Han model, a process acting on very fast time scales, turns out to have a very small effect on the predicted fluxes. On the other hand, the parameter k_r describing repair of the damaged RCIIs on a time scale of hours cannot be confidently estimated from experimental data collected over 20 minutes only. The values for τ and k_r in Table 4.3 are the mean values of the ranges reported by Han et al. (2000b). One way of determining τ experimentally would be to use FRR fluorometry that can apply flashes at microsecond intervals (Kolber and Falkowski, 1993). Likewise, a more confident estimate for k_r could be obtained by simply extending the dark phase at the end of the calibration PAM experiment, e.g., by an hour or two.

- The total number of PSUs, N, cannot be confidently estimated due to its large correlation with the total cross section σ. The value for N in Table 4.3 is based on the Emerson number of 2500 mol(chl)mol$(O_2)^{-1}$, as reported by Falkowski and Raven (1997).

- The parameter subset formed by the relative rate constant η_P, η_D, η_{qE}, η_I and the scaling factor S_F is structurally unidentifiable based on fluorescence flux measurements only, calling for fixing the value of one of these parameters. The parameter η_D representing the ratio between the rate of basal thermal decay and the rate of fluorescence can be estimated considering the probability of thermal dissipation and the probability of fluorescence for a photon absorbed by a dark adapted RCII. The value reported in Table 4.3 is the mean of those η_D values for which the resulting fluorescence quantum yields are consistent with the data by Huot and Babin (2010).

Table 4.3: *Default values of the constant parameters (left part), and estimated values with confidence intervals of the calibrated parameters (right part).*

Parameter	Value	Units
k_r	5.55×10^{-5}	s^{-1}
τ	5.50×10^{-3}	s
N	4.50×10^{-7}	$\mu E\, g_{chl}^{-1}$
η_D	5.00×10^{0}	-

Parameter	Estimate	±95% Conf. Int.	Units
ξ	5.95×10^{-2}	$\pm 6.65 \times 10^{-3}$	s^{-1}
n	2.26×10^{0}	$\pm 7.76 \times 10^{-2}$	s^{-1}
I_{qE}	8.56×10^{2}	$\pm 2.88 \times 10^{1}$	$\mu E/m^2 s$
k_d	6.41×10^{-7}	$\pm 3.38 \times 10^{-7}$	-
η_P	1.14×10^{1}	$\pm 1.60 \times 10^{-1}$	-
η_I	7.87×10^{1}	$\pm 3.94 \times 10^{1}$	-
$\overline{\eta}_{qE}$	1.98×10^{1}	$\pm 6.69 \times 10^{-1}$	-
σ	8.74×10^{-1}	$\pm 4.35 \times 10^{-2}$	$m^2 \mu E^{-1}$
S_F	1.56×10^{0}	$\pm 7.80 \times 10^{-2}$	$g_{chl}\, \mu E^{-1}\, V^{-1}$

The parameter values and 95% confidence intervals determined by gPROMS using maximum likelihood estimation are reported in the right part of Table 4.3. The corresponding fits of $F^{'}$, $F_m^{'}$ and $F_0^{'}$ against the measured fluxes are shown in Figure 4.3, both without and with the assumption $\Phi_f^A = \Phi_f^C$. Note that the predicted fluorescence fluxes are in excellent agreement with the measured fluxes, thereby providing a first confirmation that the proposed model structure captures the interplay between photoproduction, photoinhibition and photoregulation

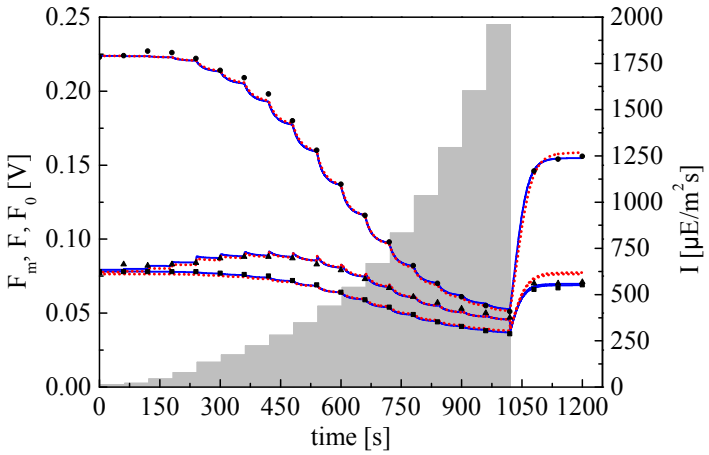


Figure 4.3: *Comparison between the predicted and measured fluorescence fluxes F'_m (circles), F'_0 (triangles) and F' (squares) in response to various actinic light levels I (grey-shaded area) for the calibration experiment. The dotted and continuous lines are obtained without and with the assumption ϕ_A equals ϕ_C, respectively.*

in a typical PAM experiment. Moreover, all the parameter estimates, but k_d and η_I, have 95% confidence interval below 10%, which is quite remarkable given the large number of estimated parameters and the apparent simplicity of the PAM protocol in Figure 4.3. Although the estimated values of k_d and η_I are found to pass the statistical t-test, the presence of a large correlation between these parameters explains their relatively poor precision. Under the assumption that $\Phi_f^A = \Phi_f^B$, or equivalently $\eta_I = \eta_P$, the 95 % confidence intervals are reduced under 10 % without significantly affecting the rest of the parameters. Nonetheless this assumption would require further testing and validation before adoption.

4.4.2 Model analysis and validation

Besides predicting the fluorescence fluxes well, the ability of the model to predict the fluorescence indexes q_L, Φ_{PS2} and q_{NPQ} is depicted in Figure 4.4, based on the expression given in the right part of Table 4.2.

The index Φ_{PS2} is predicted quite accurately by the model throughout the entire time horizon, and the smooth transition in the Φ_{PS2} profile observed once the actinic light is switched off is a consequence of qE-quenching acting directly on the quantum yield of photosynthesis in the proposed NPQ representation. Moreover, the predicted value of 0.65 for the quantum yield of photosynthesis of a dark-adapted open RCII - this value corresponds to the Φ_{PS2} at initial time here - is in excellent agreement with values widely reported in the literature (Sforza et al., 2012). The overall fitting quality of the index q_L is also satisfactory, apart from the last few experimental points during the light phase (between 800-1000 s), which are over-predicted by the model. Nonetheless, the model captures well the sharp change in q_L that occurs when the actinic light is switched off, a property that comes forward with the expression of q_L in Table 4.2 due to the fast dynamics of A and B. Finally, the accurate predictions of q_{NPQ} in

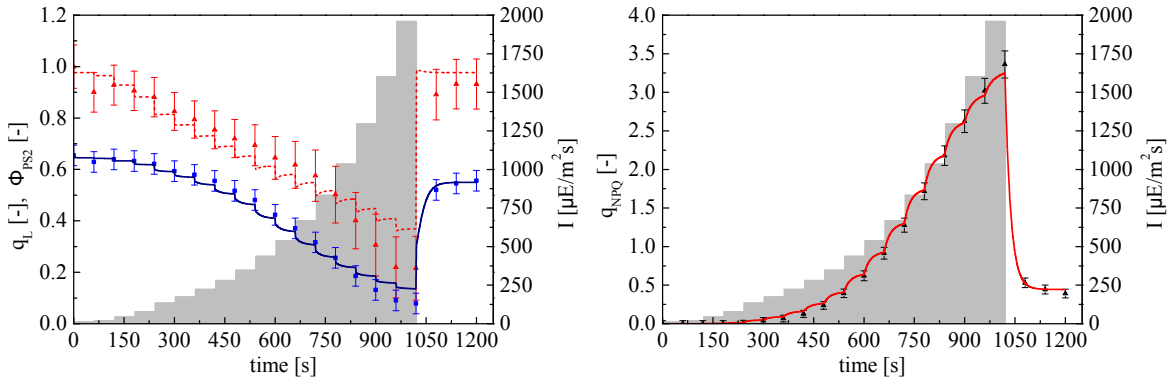


Figure 4.4: *(a) Comparison between the predicted and measured fluorescence indexes* Φ_{PS2} *(blue lines, square) and* q_L *(red lines, triangles) at various actinic light levels* I *(grey-shaded area). (b) Comparison between the predicted and measured fluorescence index* q_{NPQ}*.*

Figure 4.4(b) provide another confirmation that the NPQ regulation is captured adequately by the selected model structure.

Further validation of the model can be obtained upon analyzing the level of photoinhibition created by the continuously increasing actinic light. Specifically, the main plot on Figure 4.5 shows a comparison between the fraction C of damaged RCIIs predicted by the full calibrated model and the same fraction given by (4.15). We recall that the later uses the available fluorescence flux measurements in combination with the predicted qE-activity level α, but does not rely on the Han model at all. These two damage fractions are found to be in good agreement, especially when considering the error bars and the red envelope of predictions computed from the 95% confidence intervals of the calibrated parameters in both cases. These rather large errors - between 0.02-0.07 at the end of the light phase - are caused by the rather large confidence intervals for the parameters k_d and η_I in this case. For comparison purposes, the smaller plot on Figure 4.5 shows the predictions of the alternative model of qE quenching, whereby the qE activity variable α modifies the absorption cross section σ in (4.1) instead of the quantum yield of fluorescence Φ_f. As well as the large discrepancy between both predictions of the damage level, it is the unusually large damage level (up to 60% after 1000 s) along with the fast repair rate that clearly invalidate this alternative qE-quenching representation.

The foregoing results suggest that the proposed fluorescence model is capable of quantitative predictions of the state of the photosynthetic apparatus under varying light conditions. To confirm it, we carry out a validation experiment for an (unusually) challenging PAM experiment, as shown in grey-shaded area on Figure 4.6. The corresponding model predictions, based on the default and calibrated model parameters in Table 4.3, are compared to the actual flux measurements in Figure 4.6. Although calibrated for a quite different and somewhat simpler PAM protocol, the calibrated model remains capable of reliable quantitative predictions of the fluorescence fluxes. Deviations are observed in various parts of the response flux profiles, which are possibly due to effects and processes not accounted for in the proposed model, yet these

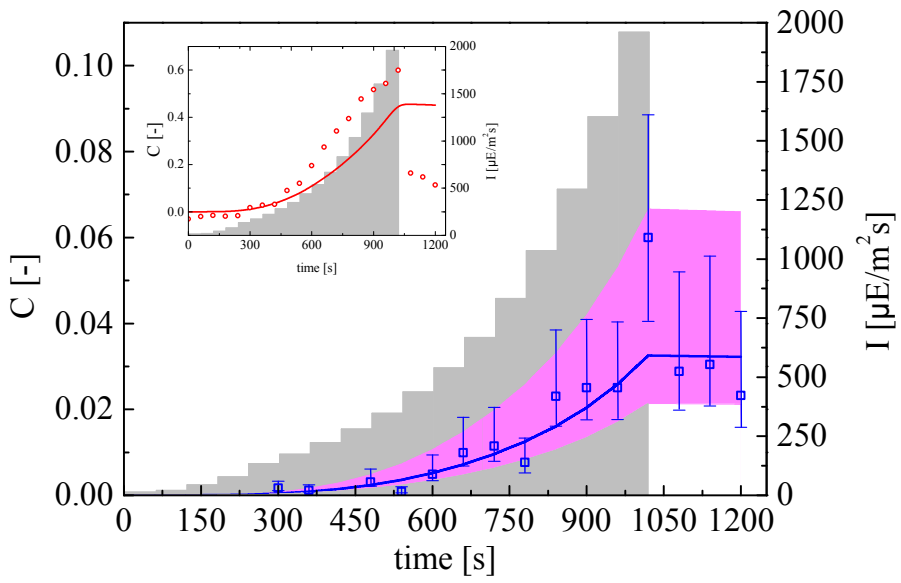


Figure 4.5: *Large plot: Comparison between the fraction C of damaged RCIIs predicted by the full calibrated model (blue lines) and by the expression* (4.15) *(blue squares) at various actinic light levels I (grey-shaded area). Small plot: Similar comparison for the variable σ model alternative (red lines and circles).*

deviation remain small, within 10-20%. We also note that such extreme variations of the light conditions, however useful in a model validation context, are unlikely to be found in a practical microalgae culture systems.

4.5 Conclusions

In this chapter we presented a mathematical representation of key photosynthetic processes acting on time scales up to an hour and triggered by varying light conditions, which are typical in PAM experiments. The dynamic fluorescence model relies on the combination of fast photosynthetic mechanisms with slower photoprotective mechanisms in order to yield a light-dependent expression of the quantum yield of photosynthesis. Despite comprising a total of 13 parameters, a careful calibration and subsequent validation against multiple experimental data sets shows that the model is capable of quantitative predictions of the state of the photosynthetic apparatus in terms of its open, closed and damaged reaction center. This makes it the first model of its kind capable of reliable predictions of the levels of photoinhibition and NPQ activity, while retaining a low complexity and a small dimensionality.

Such generic capability to predict the development of photoinhibition and photoregulation, yet without the need for dedicated experiments (Ruban and Murchie, 2012), addresses a long-standing challenge in the modelling of photosynthetic productivity and holds much promise in regard of future applications. By design, the fluorescence model is indeed capable of simulating experimental protocols used for the determination of PI-response curves, yet avoiding the usual - and somewhat problematic - static growth assumption. This opens the possibility

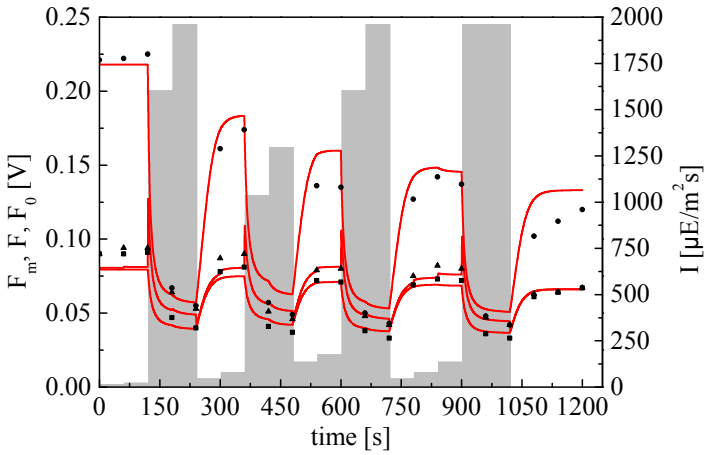


Figure 4.6: *Comparison between the predicted and measured fluorescence fluxes F'_m (circles), F'_0 (triangles) and F' (squares) in response to various actinic light levels I (grey-shaded area) for the validation experiment.*

for a cross-validation framework, whereby both fluorescence and classical growth experiments could be used for model validation purposes. This issue will be addressed in Chapter 5. In combination with dedicated PAM experiments, there is also hope that the model could serve as a platform for unveiling previously hidden information concerning the operation of the photosynthetic apparatus. Because PAM experiments are both precise and fast, a full validation of the model appears tractable in this context, especially if model-based experimental design is used for testing the model structure further, e.g., through the determination of information-rich PAM protocols. A MBDoE approach will be presented in Chapter 6.

Incorporating photoacclimation processes is currently investigated as part of future work, and preliminary results are presented in Chapter 5, in order to widen the applicability of the model, such as predicting the evolution of a microalgae culture over time periods of several days or even weeks. Eventually, the vision is to integrate a fully validated model of photosynthesis within first-principle models describing the flow and light attenuation in large-scale microalgae culture systems as a means to guide their design and operations.

Chapter 5

A framework for the dynamic modelling of PI curves in microalgae

In this chapter[1] we present an extension of the model discussed in Chapter 4 able to predict the photosynthesis rate under dynamic light conditions. Photoproduction, photoinhibition and photoregulation are well represented once the model is validated against experimental data of chlorophyll fluorescence for the microalga *Nannochloropsis gaditana*. Photoacclimation effects are incorporated using fluorescence data for three photoacclimation states. The model shows a very good agreement with the available experimental data and the values of the calibrated parameters are deemed valid from a biological viewpoint. The predicted PI exhibits reasonable trends regarding their initial slope, maximum photosynthesis rate and photoinhibitory effects. Moreover, we will show that different PI curve characteristics arise as a result of different experimental protocols thus underlying the importance of the accurate description of the protocols used to derive the experimental data. Finally, a preliminary calibration against experimental data from the literature is attempted to show the descriptive capability of the model.

5.1 Motivation

Photosynthesis-Irradiance (PI) response curves constitute a powerful tool in characterizing the photosynthetic properties of microalgae. Accurate representation of such curves in a mechanistic manner has been an important challenge for many years, since both the effects of light and nutrients can be captured and optimal productivities of large scale production systems can be inferred (Bernard, 2011). The most usual approach in representing PI curves is in the form of static relationship between the light irradiance and the photosynthesis rate, the latter is often expressed in terms of oxygen evolution or CO_2 fixation as measured using of oxygen electrodes or radioactive carbon tracing (Jassby and Platt, 1976). Nonetheless, the

[1]Part of this chapter has been published in Bernardi et al. (2015a)

inherently dynamic nature of microalgae cultures, together with the experimental protocol used to obtain such PI curves, introduces a great level of uncertainty in the actual measurements and also raise legitimate questions as to the validity of the underlying biological hypothesis of a static model. Moreover, the experimental procedure used to obtain PI curves is time consuming and are carried out in conditions that are very different from the growing conditions in the photobioreactors (e.g. no mixing and no CO_2 excess).

The main objective of this chapter is to extend the model presented in Chapter 4 in order to predict photosynthesis rate for different acclimation states. The model was calibrated and validated using data from PAM fluorometer showing very good agreement between the experimental data and the model predictions. The model is able to represent in sufficiently accurate way all the dynamic processes involved in photosynthesis. Once validated the model can be used to predict PI curves from fast and accurate fluorescence measurements.

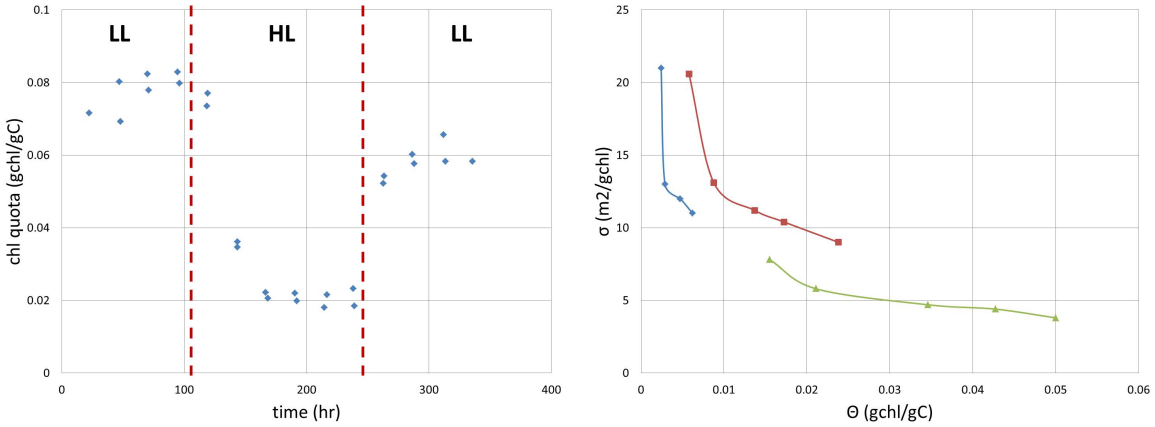
In particular, dynamic simulations of PI curves allows to evaluate the effect of the experimental protocol in the PI curves characteristic. Despite the rather common hypothesis from the literature that PI curves are independent from the protocol used to obtain them, we will show the importance of taking into account the exact experimental protocol to prevent misleading conclusions or wrong parameter estimations. The effect of the initial amount of damaged PSUs will be analysed as well. Finally, a preliminary calibration against experimental data from the literature is attempted to show the descriptive capability of the model. The results will show good agreement between the experiment and the model prediction even if the uncertainty regarding the experimental protocol still represent a limitation in the analysis.

5.2 Dynamic modelling of Photosynthesis-Irradiance curves

The fluorescence model in Chapter 4 accounts for photoproduction, photoregulation and photoinhibition. To complete the model, the photoacclimation process has to be included. In the following the main effect of photoacclimation will be discussed. Next an expression for the photosynthesis rate will be introduced.

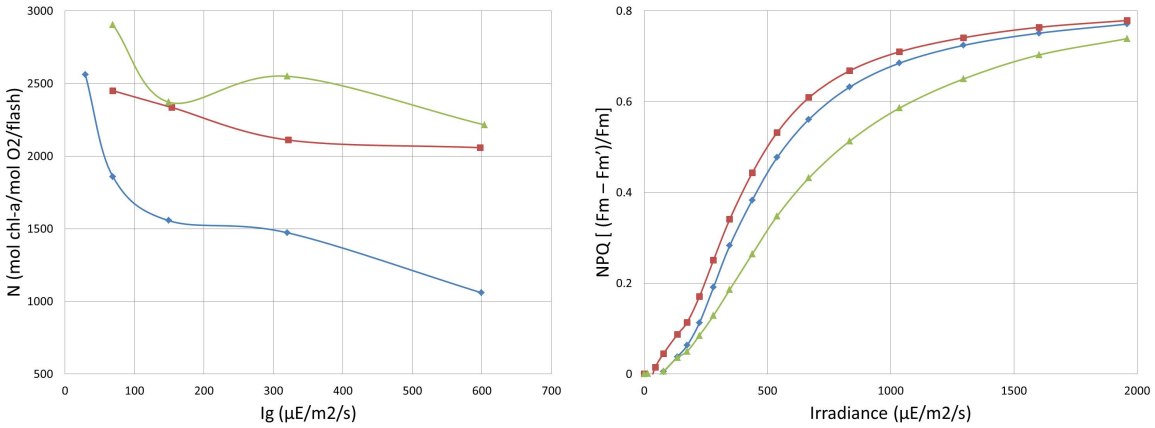
5.2.1 Photoacclimation extension

Photoacclimation refers to physiological adaptations of the photosynthetic apparatus that enables microalgae to optimally grow under long term variations in the light intensity (MacIntyre et al., 2002). The most profound effect of photoacclimation is a decrease in the chlorophyll content, when irradiance is increased and vice-versa (Fisher et al., 1996; Dubinsky and Stambler, 2009). Figure 5.1(a) reports the experimental data from Anning et al. (2000) demonstrating a 4-fold variation in the chlorophyll-a composition when the culture is transiting from low light (LL) to high light (HL).



(a) (b) effect of C_0

Figure 5.1: *(a) Experimental data of chlorophyll content by Anning et al. (2000) (b) Experimental data of total cross section by Falkowski et al. (1985).*



(a) (b) effect of C_0

Figure 5.2: *(a) Experimental data by Falkowski et al. (1985) demonstrating the variation of parameter N (b) Experimental data by Sforza et al. (2012) demonstrating the different activity of NPQ for different acclimation states.*

The chlorophyll variations affect the photosynthetic apparatus in several ways:

- The total cross section, σ decreases as chlorophyll increases due to shading effects. Literature data from Falkowski et al. (1985) are reported in Figure 5.1(b);
- the number of chlorophyll associated with each PSU, i.e. the reverse of parameter N decreases with the growth irradiance. Literature data from Falkowski et al. (1985) are reported in Figure 5.2(a).

Finally the variation of pigment composition affects the activity of photoregulation (Anning et al., 2000) as we can observe from Figure 5.2(b). Accordingly, microalgae samples, acclimated at different light intensities show different fluorescence responses under identical PAM protocols. Our data comprise three standard PAM experiments for three samples acclimated to three

different growing irradiances, I_g. As the experiments considered to calibrate the model are only 20 minutes long we can confidently assume that the acclimation dependent parameters are constants within the same experiment.

5.2.2 Photosynthesis rate

The photosynthesis rate expression comes forward as (Falkowski et al., 1994):

$$P = \sigma \, \Phi \, I \tag{5.1}$$

where Φ is the oxygen photosynthesis quantum yield [$\mathrm{mol_{O_2}}/\mu$E]. The aforementioned units, give dimensions for P in [$\mathrm{mol_{O_2}}/\mathrm{g_{chl}s}$], which by definition is the chlorophyll specific photosynthesis rate, in terms of O_2 production. The dynamic model of fluorescence predicts the value of the realised quantum yield of photosynthesis, Φ_{PS2} in units of [$\mathrm{mol\,e^-}/\mu$E], which is known from the literature to be closely related to Φ (Suggett et al., 2003). We only need to introduce a stoichiometric coefficient that aligns the electrons delivered in RCII to the O_2 produced there. If we consider the water dissociation reaction in PSII we can derive a theoretical minimum value of 4 $\mathrm{mol\,e^-/mol\,O_2}$ as conversion factor.

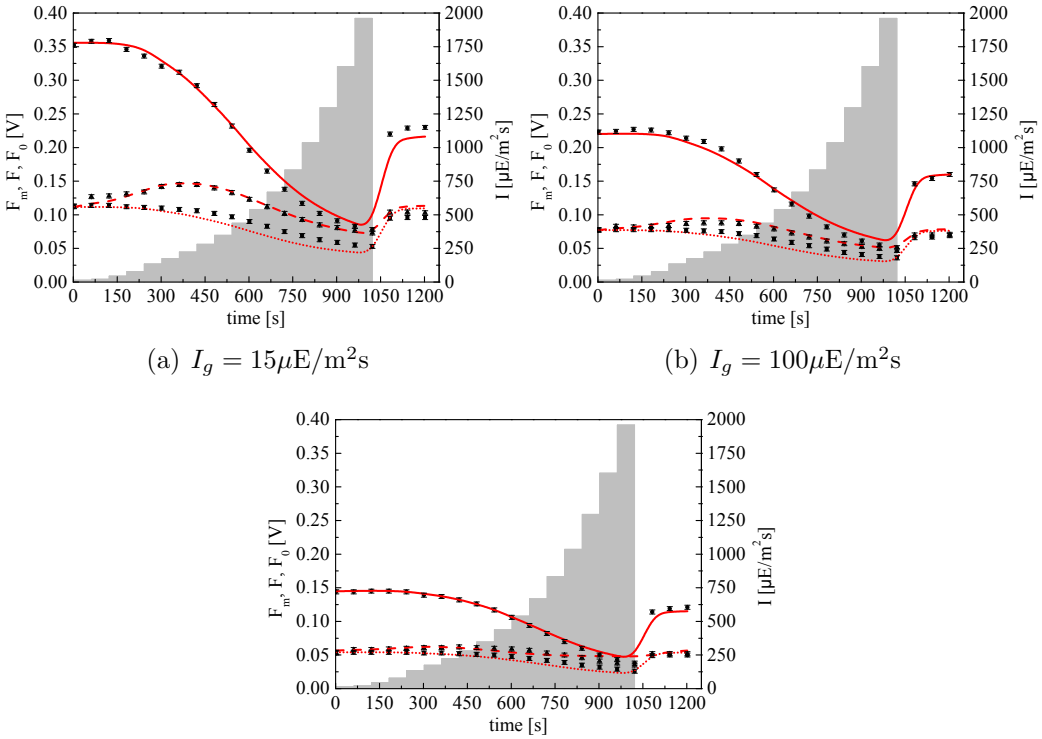
5.3 Results and discussion

The model has been calibrated using data from three independent fluorescence experiments performed with samples acclimated to different I_g (15, 100 and 625 $\mu\mathrm{E\,m^{-2}\,s^{-1}}$). We will not propose empirical equation to describe photoacclimation as the available data are not sufficient to discriminate between different modelling hypotheses and beyond the scope of the work. For this reason we will use different parametrisation to account for photoacclimation: each acclimation state will have unique values for the parameters σ,S_F, N, I_{qE} and η_D. The remaining parameters take the same values for all three model replicates.

Table 5.1 reports the values of the acclimation dependent parameters. The fitting results for the fluorescence fluxes are presented in Figure 5.3, which shows a good agreement between the model and the experimental data.

Table 5.1: *Parameter values for different photoacclimation states.*

Parameter	15 μE/m^{2}s	100 μE/m^{2}s	625 μE/m^{2}s
I_{qE}	796	847	1076
σ	7.86×10^{-1}	8.74×10^{-1}	9.44×10^{-1}
η_D	4.80	5.00	5.75
N	3.47×10^{-7}	4.50×10^{-7}	6.62×10^{-7}
S_F	2.59	1.56	1.05



(a) $I_g = 15\mu\text{E/m}^2\text{s}$ (b) $I_g = 100\mu\text{E/m}^2\text{s}$

(c) $I_g = 625\mu\text{E/m}^2\text{s}$

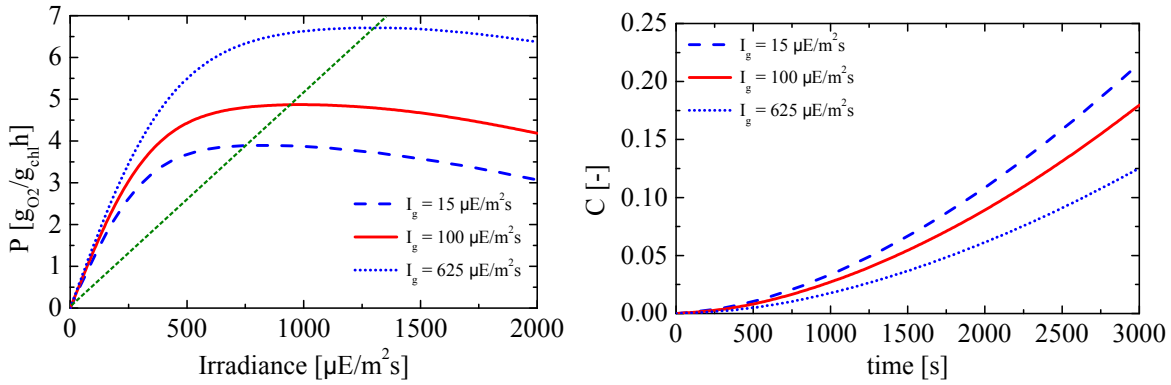
Figure 5.3: *Fitting results of the fluorescence fluxes for 3 photoacclimation states. Upper, middle and lower plots correspond to 625, 100 and 15* $\mu\text{E/m}^2\text{s}$.

Regarding the parameter estimates in Table 5.1 we can observe that they are consistent with the biological meaning following the trends mentioned in Section 5.2.1. In particular: S_F, which indirectly represents the chlorophyll content, shows a decreasing trend with growth irradiance; σ is inversely related to S_F due to the packaging effect; N is decreasing with increasing *chl*; η_D is increasing as growth irradiance increases and that can be linked to the accumulation of photoprotective pigments; the relation of I_{qE} and growth irradiance is almost linear, and it is reasonable that microalgae acclimated at higher I_g need to activate the NPQ at higher irradiances.

5.3.1 Dynamic simulation of PI curves

In order to construct the PI curves a light protocol with a step profile for the light has been considered. The light was kept constant for 5 minutes and then increased by 200 $\mu\text{E/m}^2\text{s}$ ranging from 0 to 2000 $\mu\text{E/m}^2\text{s}$. The value at the end of each light phase was then plotted against the irradiance to determine the PI curve. The initial fraction of C was assumed to be zero based on the experimental data used for model calibration.

In Figure 5.4 we can observe that the predicted PI curves are very similar to published data. The initial slope is constant, and the maximum photosynthesis rate is greater for higher acclimated cells, two phenomena widely reported in the literature (Anning et al., 2000).



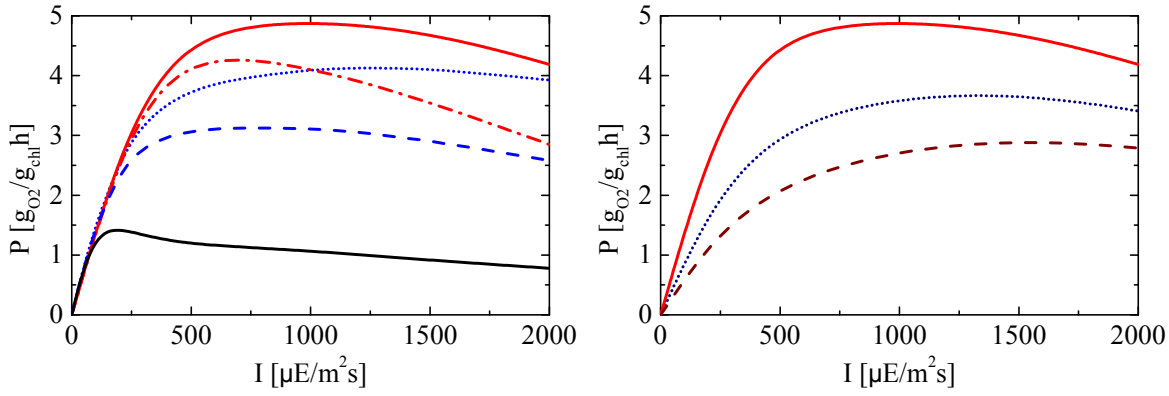
(a) PI curves (b) C profile

Figure 5.4: *PI curves (a) and evolution of damaged RCIIs (b) for 3 photoacclimation states. Dotted, dashed and continuous lines correspond to 625, 100 and 15* μE/m^{2}s.

Moreover the optimum light intensities and the maxima of photosynthesis rate for different acclimation states are linearly related as reported in the literature (Hartmann et al., 2013). Finally, the damaged RCII profiles, reported in Figure 5.4 show that cells acclimated at lower light (hence with higher chlorophyll content) are more prone to inhibition, which is also consistent with biological understanding.

With respect to the sample acclimated to 100 μE/m^{2}swe now want to assess the effect of the light protocol. In the following we will compare two different experimental set up. In the first experimental set up, indicated as Type A, we consider that to determine an experimental point a sample of the culture has to be kept at constant light for a certain amount of time, called incubation time, before measuring the value of P. The second experimental set up, indicated as Type B, considers to have only one sample exposed to a varying light intensity. The light will follow a step profile with the light increasing from zero to a maximum value and P will be measured at the end of each constant light step. Note that in the literature PI curves data are, in majority, reported without the experimental protocol that was used to obtain them. We will show how dynamic modelling plays a crucial role and the experimental protocol can lead to different PI curves behaviour.

Figure 5.5(a) compares the Type A and Type B experiments. For each experimental set up two alternatives are considered regarding the duration of the experiment. The blue (dotted and the dashed) line refers to Type A experiment. Ten samples are assumed to be exposed independently to ten different light intensities ranging from 0 to 2000 μE/m^{2}s. The incubation time has been set to 1800 s for the dotted line and to 3600 s for the dashed line. The red (continuous and the dot-dash) line refers to Type B experiment. The irradiance is assumed to increase from 0 to 2000 μE/m^{2}sfollowing a step profile with ten constant light steps. Each step lasts 300 s for the continuous line and 600 s for the dash-dot line; at the end of each step P is measured and the light is increased by 200 μE/m^{2}s. The continuous black line represents the steady state PI profile predicted by the model. Figure 5.5(b) investigates the effect of the



(a) effect of the experiment (b) effect of C_0

Figure 5.5: *(a) PI curves for a sample acclimated at 100 μE/m^{2}sand different light protocols: dotted line considers a Type A experiment with an incubation time of 1800 s and dashed line considers a Type A experiment with an incubation time of 3600 s; continuous red line consider a Type B experiment with constant light steps of 300 s; dash-dot line considers Type B experiment with steps of 600 s constant light. Black continuous line represent the steady state PI profile predicted by the model. (b) PI curves obtained for a Type B experiment with 300 s constant light step and different initial conditions for damaged reaction centres.*

initial condition of damage in the predicted profile if a Type B experiment is considered (Type A results lead to equivalent conclusions). The three curves of Figure 5.5(b) consider a step duration of 300 s. The continuous line is the same as in Figure 5.5(a), while the dashed line considers an initial value of C (i.e. C_0) equal to 0.1 whereas the dotted line considers C_0 equal to 0.2.

From Figure 5.5(a) we can observe how both the experimental set up and duration affect the PI curves. The type of experiment affects the shape of the PI curves, while the initial slope remain constant. Note that the experiment duration affects the behaviour of PI curves at high light intensities. In fact, a longer experiment will cause an higher amount of photoinhibition at high light intensities, thus leading to diminished photosynthetic production. Moreover, we can observe that the steady state PI photosynthesis rate profile is significantly lower than the dynamic profiles. In practical terms, this means that if someone uses the dynamic profiles as an estimate of the steady state photosynthesis rate he can drastically overestimate the culture performances. From Figure 5.5(b) we can observe that, if at the beginning of the protocol the sample has a certain amount of inhibited reaction centres, we have a reduction both on the initial slope and on the maximum photosynthesis rate. In view of the above, we can conclude that it is necessary to consider a dynamic model and that the exact experimental protocol used to obtain the PI curves should be taken into account, if misleading conclusions or wrong parameter estimation are to be avoided. Also the initial condition of damage needs to be evaluated in order to correctly compare different PI curves. This is particularly important if we consider cells acclimated at high light conditions, where an initial damage is likely to be present as consequence of the stressful growing environment.

5.4 Coupling PI curves with fluorescence data for model calibration

In the previous section we showed how dynamic modelling have to be taken into account when considering PI curves experiments. In fact, since different PI curve characteristics can be obtained varying the experimental protocol used to obtain them, only a dynamic model that implement the exact experimental protocol can represent the system in a reliable way.

In this section we want to assess the potential advantages of coupling PI curves measurements with fluorescence data. In particular, we want to understand if the effect on the quality of parameter estimation, in terms of the estimation accuracy and number of estimated parameters, is significant. Moreover we also want to quantify the error related to a wrong hypothesis regarding the experimental protocol. In particular, we want to answer to the following questions: (i) is it possible to discriminate between alternative hypothesis regarding the experimental protocol using the fluorescence model? and (ii) which are the parameters affected by a wrong hypothesis regarding the experimental protocol?

In order to answer the above mentioned questions we will use an in silico experiment. In fact, using a simulated experiment we know the correct values of the parameters and the exact experimental protocol used to obtain the data; therefore is an easy task to precisely quantify the error induced by a wrong hypothesis on the protocol.

Finally, a PI curve from the literature (Gentile and Blanch, 2001) will be used to show the model behaviour against real experimental data. Unfortunately, this will not lead to a proper validation, since in Gentile and Blanch (2001) the acclimation state of the culture was different from the acclimation state of the samples used to perform the fluorescence experiments and the experimental protocol to obtain the photosynthesis rate was not clearly specified. For this reason the estimates obtained can not be confidently used and new data will be required to have an accurate estimation of parameter N.

5.4.1 Case study: in silico PI curves experiment

The continuous curve of Figure 5.4(a) is considered as an in silico experiment to be used in the calibration set, along with the PAM experiment. The utilisation of an in silico experiment allow to analyse the response of the model under the hypothesis that the model perfectly represent the data. Moreover, as we know the light protocol and the correct values of the parameters we can analyse the effect of assuming the wrong experimental protocol in the parameter estimation.

If we estimate the model parameters using the correct light protocol an accurate parameter estimation can be achieved. Moreover, parameter N, that has to be fixed if we only consider a PAM experiment, can be confidently estimated.

In Table 5.2 we can observe that the parameter N is very precisely identified with a confidence interval 95% smaller that 10% of the estimated value. Moreover all the parameters

Table 5.2: *Parameter values along with 95% confidence interval and t-values if a PI curve measurement is considered in the calibration set along with the PAM experiment used in Chapter 4. The parameter values are normalised by the values obtained from calibration in Chapter 4. Reference t-value is 1.67.*

Par	Norm val	95% c.i.	t-val 95%
ξ	1.00	1.11×10^{-1}	8.98
I_{qE}	1.00	3.36×10^{-2}	29.81
k_d	1.02	4.93×10^{-1}	2.06
k_r	1.00	-	-
N	1.00	8.34×10^{-2}	12.03
n	1.00	3.42×10^{-2}	29.23
η_I	0.98	4.53×10^{-1}	2.17
η_{qE}	1.00	3.36×10^{-2}	29.71
η_D	1.00	-	-
η_P	1.00	1.37×10^{-2}	72.69
S_F	1.00	8.02×10^{-2}	12.41
σ	1.00	8.12×10^{-2}	12.36
τ	1.00	-	-

are estimated correctly (the values are normalised by the "real" value used to simulate the in silico experiment). We can therefore conclude that, if the protocol used to obtain the PI curve is known, the information contained in the experiment allows the accurate estimation of parameter N, which otherwise is unidentifiable.

We now want to study the error that a wrong assumption about the experimental protocol can introduce in the parameter estimation. The in silico experiment used has been obtained with a Type B experiment with 300 s constant light step. We will now considers two alternative scenarios: in Scenario 1 we assume a Type A experiment with 3600 s incubation time; in Scenario 2 we assume the correct experiment type but the wrong constant light step duration (600 s).

In Figure 5.6 the profiles for the two alternative scenarios are reported. The dotted line refers to Scenario 1, while the continuous line refers to Scenario 2.

We can observe that in Scenario 1 (wrong experiment type) the model can represent the data with good accuracy, while in Scenario 2 (correct experiment type but different experiment duration) the fitting quality is lower but still acceptable. This result points out that is not possible to discriminate between alternative protocols based on model simulation and therefore a lack of knowledge about the experimental protocol is a systematic error in the parameter estimation difficult to detect.

It is important to underline that, the fact that the model is able to fit the data even with a wrong hypothesis about the experimental protocol does not mean that the model is unidentifiable or overparametrised. In fact, a wrong assumption about the experimental protocol affects the parameter estimation results, resulting in a wrong estimation. In Table 5.3

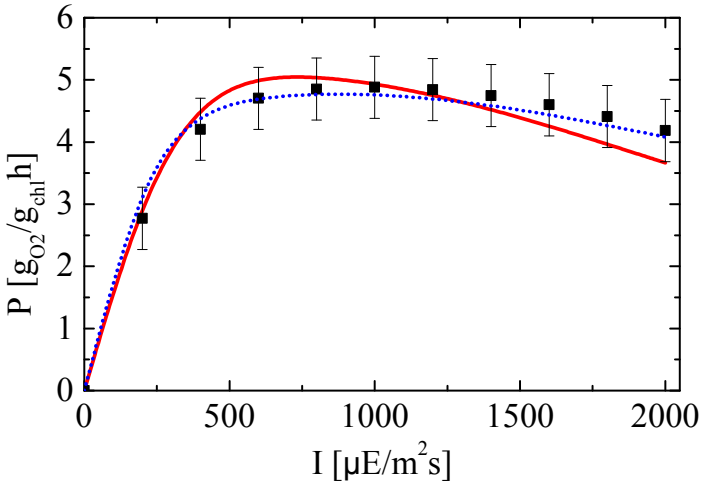


Figure 5.6: *The squares represent the in silico experiment obtained by model simulation and assuming a Type B experiment with 300 s constant light steps. The continuous line is the predicted profile if a calibration is attempted considering a Type B experiment with 600 s constant light steps; the dotted line consider the predicted profiles if a calibration is attempted considering a Type A experiment with 3600 s incubation time. Table 5.2 and Table 5.3 summarise the results of the two parameter estimations.*

are reported the estimates along with the t-values obtained in the aforementioned scenarios.

We can observe in Table 5.3 that both scenarios lead to wrong parameter estimations. In particular, the subset of model parameters that are affected by the wrong hypothesis about the protocol comprises: σ, N, S_F, k_d and η_I. In Scenario 1 the estimation of parameters σ, N and S_F are different from the nominal values. Also parameters k_d and η_I are slightly different from the nominal values but the relatively large confidence intervals include the correct values. In Scenario 2, where the model showed larger discrepancy with the in silico data, parameters σ, N and S_F are again wrongly estimated. Moreover, the model mismatch introduced with the wrong assumption regarding the experimental protocol, lead to a decreased accuracy of parameter estimation and in particular parameters k_d and η_I fails the t-test.

To summarise, a calibration has been attempted using an in silico experiment along with a standard fluorescence experiment. The information contained in the photosynthesis rate profile lead to a precise parameter estimation and to the identification of one extra parameter, N. In order to assess the effect of a wrong assumption regarding the protocol used to obtain the PI curve two additional scenarios have been considered. The results show that the model is able to represent with good accuracy the data, even if a wrong protocol is assumed, but the resulting parameter estimations are not consistent.

5.4.2 Case study: calibration with literature data

A preliminary calibration based on real PI curves data has been carried out considering data from the literature. The data by Gentile and Blanch (2001) consider a sample of Nannochloropsis Gaditana acclimated at 200 μE/m^{2}s. In the paper the authors both describe Type A and Type B experiments but it is not clear which protocol they have used to obtained

Table 5.3: *Parameter estimates, confidence intervals 95%and and t-values 95% considering the PI curve in silico experiment in the calibration set if a wrong light protocol is assumed. Reference t-value is 1.67. Statistically unsatisfactory estimates are indicated by (*).*

Par	Norm val	95% c.i. Scenario 1	t-val 95%	Norm val	95% c.i. Scenario 2	t-val 95%
ξ	1.00	1.10×10^{-1}	9.11	0.99	1.08×10^{-1}	9.18
I_{qE}	1.00	3.34×10^{-2}	29.89	1.00	3.35×10^{-2}	29.87
k_d	0.79	4.48×10^{-1}	1.77	0.67	4.54×10^{-1}	**1.47 ***
k_r	1.00	-	-	1.00	-	-
N	1.48	1.59×10^{-1}	9.31	1.17	1.06×10^{-1}	11.00
n	1.00	3.41×10^{-2}	29.3	1.00	3.42×10^{-2}	29.22
η_I	1.24	6.75×10^{-1}	1.84	1.48	9.76×10^{-1}	**1.56 ***
η_{qE}	1.00	3.35×10^{-2}	29.78	1.00	3.35×10^{-2}	29.76
η_D	1.00	-	-	1.00	-	-
η_P	1.00	1.34×10^{-2}	74.0	0.99	1.35×10^{-2}	73.78
S_F	0.68	7.07×10^{-2}	9.55	0.86	7.52×10^{-2}	11.38
σ	1.48	1.56×10^{-1}	9.47	1.16	1.03×10^{-1}	11.31
τ	1.00	-	-	1.00	-	-

the PI curve data. For this reason we will consider two alternative calibrations, assuming either Type A or Type B experiments. The PI curve will be used to calibrate the model along with the PAM experiment in Figure 5.3(b). The aim is to assess if the model is able to reproduce accurately literature data and to test the practical identifiability of parameter N. One of the main limitation of the analysis is that the acclimation state of the sample used to perform the fluorescence experiment is different from the sample used by Gentile and Blanch (2001). However, the difference in the growth irradiances is not very high and the microalgae are expected to have similar characteristics. Another limitation is related to the lack of information regarding the experimental protocol.

Since the photosynthesis rate is zero at 120 μE/m^{2}sa constant respiration rate, R, has been assumed. Accordingly, Equation 5.1 has to be modified in order to account for respiration:

$$P = \sigma \, \Phi \, I \, - R \tag{5.2}$$

In the following the model parameters will be normalised by the values obtained during the calibration in Chapter 4 except for R. Note that the respiration rate has not been considered in the previous sections because the focus was not on the absolute values but only on the shape of the PI curves, and therefore subtracting a constant term did not affect the results. On the contrary, when real experimental data are considered the respiration rate is necessary, as it represents the oxygen consumption in the dark, that is usually greater than zero.

From Figure 5.7 we can observe that both Type A and Type B experiment can explain the experimental data. Type A experiment seem to be more accurate in capturing the trend of

Table 5.4: *Parameter estimates, confidence intervals 95%and and t-values 95% considering a PI curve from the literature. The results considered a Type A experiment with 3600 s of incubation time or a Type B experiment with a 300 s step duration. Reference t-value is 1.67. Statistically unsatisfactory estimates are indicated by (*).*

Par	Norm val	95% c.i.	t-val 95%	Norm val	95% c.i.	t-val 95%
	Type A experiment			Type B experiment		
k_d	0.11	4.34×10^{-1}	**0.25 ***	0.80	4.80×10^{-1}	1.68
N	2.41	3.20×10^{-1}	7.54	2.08	2.52×10^{-1}	8.28
R	3.81	7.90×10^{-1}	4.81	3.57	7.39×10^{-1}	4.84
η_I	8.97	3.60×10^{1}	**0.25 ***	1.24	7.08×10^{-1}	1.75
S_F	0.42	5.46×10^{-2}	7.64	0.48	5.81×10^{-2}	8.31
σ	2.39	3.14×10^{-1}	7.59	2.07	2.50×10^{-1}	8.27

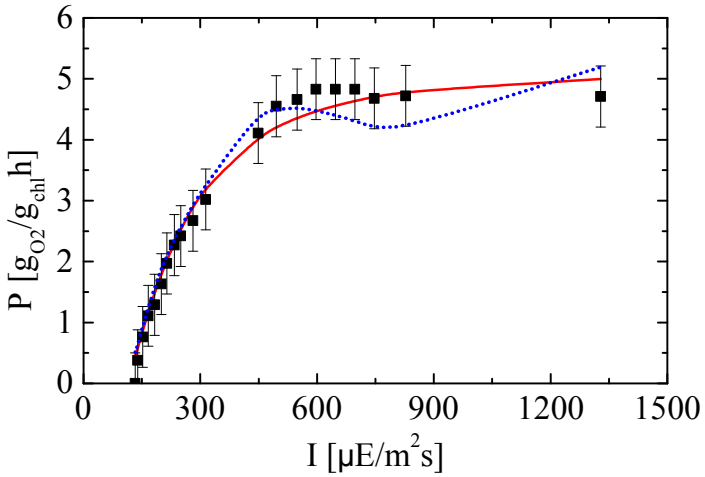


Figure 5.7: *The squares represent PI curve measured by Gentile and Blanch (2001). The continuous line is the predicted profile if a Type A experiment with 1800 s incubation time is assumed. The dotted line is the predicted profile if a Type B experiment with 300 s constant light steps is assumed. Table 5.4 summarise the results of the two parameter estimations.*

photosynthesis rate against the irradiance but also Type B experiment lead to a satisfactory prediction. As analysed in the previous section the assumption related to the experimental protocol affects the values of a subset of the model parameter. The parameter estimation results on the data by Gentile and Blanch (2001) confirm the results obtained with the in silico experiment, thus, in Table 5.4 we only reports the parameters that are affected by the introduction of the new experiment (i.e. σ, N, S_F, k_d and η_I), as the other parameters are the same as in Table 5.2.

We can observe that if a type A experiment is assumed the parameters k_d and η_I fail the t-test, while if a type B experiment is assumed the parameter estimation is more accurate. Moreover, the parameter values obtained by the two calibrations are very similar, except for parameter k_d (this is due to the fact that most of the experimental points are obtained at low-medium irradiances where the difference between the profiles predicted for different experimental set-up is lower). The main issue related to the Type B experimental set up is

concerned with the unevenly distributed light intensities for the step profiles. This results in a more complex profile in the second part of the experiment, which does not seem to be fully representative of the available data. To conclude, although the model fit the PI curve experiment, the estimates can not be confidently used due to the uncertainty regarding the experimental protocol and the different acclimation state of the sample used for PI curves determination with respect to the sample used in the PAM experiment.

5.5 Conclusion

The chlorophyll fluorescence model presented in Chapter 4 has been extended to be able to predict photosynthesis rate. The semi-mechanistic representation of the photoproduction, photoregulation and photoinhibition phenomena of the fluorescence model provide an accurate representation of the biological processes occurring in time-scales for seconds to hours, whereas photoacclimation is accounted for by varying a subset of parameters in the model. The model is able to represent the available experimental fluorescence data for three different acclimation states. Moreover the model is able to predict photosynthesis-irradiance response curves based only on fluorescence measurements. Once validated the model will allow to predict PI curves using fast and reliable fluorescence measurement, instead of performing a time consuming and inaccurate PI curve protocol.

Simulation results underline that is extremely important for the utilization of PI in estimation of biomass productivity to specify the exact protocol followed to measure PI curves and the initial condition of the photosynthetic apparatus. Moreover, the effect of coupling fluorescence and PI measurements has been assessed. An in silico experiment has been used to analyse the error in parameter estimation if the wrong experimental protocol is assumed and the practical identifiability of the model, under the assumption that the model perfectly represents the data. Finally, a preliminary calibration against literature data has been carried out to show that the model is capable of representing experimental PI curves with good accuracy and confirm the practical identifiability results of the theoretical case study.

Chapter 6

MBDoE approach for information rich PAM experiments

This chapter[1] aims to improve the accuracy of parameter estimation of the fluorescence model presented in Chapter 4. Model based design of experiments (MBDoE) techniques are at the core of the proposed procedure. Local sensitivity analysis as well as a-posteriori identifiability testing will be carried out to underline possible problems on model identification. As discussed in Chapter 4 four out of thirteen parameter have to be fixed and two of the estimated parameters have large confidence intervals. In this chapter a new type of fluorescence measurement has been used to have an estimation of the dynamic of photoproduction, that cannot be investigated by the PAM fluorometry due to its very fast time scale. Next, experiments with long recovery phases have been considered to estimate the recovery rate of damaged PSUs. Finally, a sequential MBDoE approach will be applied to develop information rich PAM experiments.

6.1 Motivation

In Chapter 4 a novel dynamic model of chlorophyll fluorescence has been presented and discussed. The fluorescence model represents with excellent accuracy the experimental data and is able of quantitative prediction of fluorescence flux of a complex validation experiment. The model can be used also for the estimation of damaged PSU during a PAM experiment. However, the model showed some identification issues even if a very small experimental error has been assumed: four out of thirteen parameters needs to be fixed to literature values and two parameters has large confidence intervals. In particular, the large confidence intervals of parameter k_d lead to a large error bars in the estimation of inhibited PSU. In this Chapter we will assume a more realistic variance model where a constant term (equal to 0.0015), related to the sensitivity of the photomultiplier that measure the fluorescence, is added to a relative term (1% of the fluorescence intensity), depending upon the intensity of the fluorescence flux.

[1]Part of this chapter has been published in Bernardi et al. (2015b)

The utilization of a more realistic variance model leads to a worsening of parameter estimation thus advocating for a design of experiments.

The main issues we want to address in this Chapter are: (i) to introduce a new type of measurement, called antenna size measurements, for the estimation of parameter τ; (ii) to reduce the confidence intervals of parameters k_d and η_I through a MBDoE approach; and, (iii) to estimate parameter k_r considering experiments with a long recovery phase in the calibration set. The long recovery phase experiments will underline a model incapability at representing NPQ dynamics and a more detailed description of the photoregulation will be necessary to describe the data.

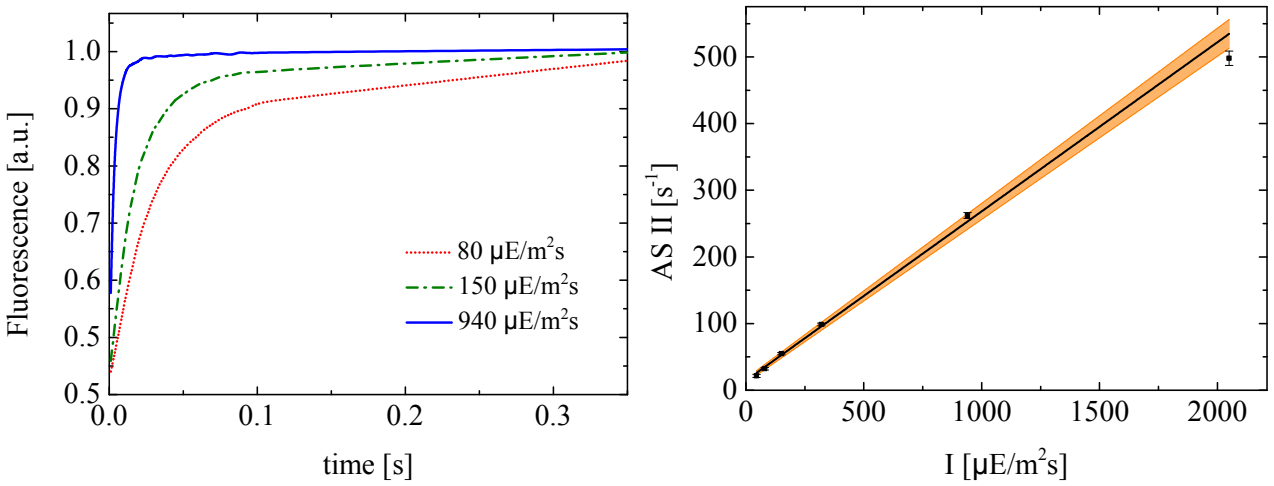
After introducing the antenna size measurements for the estimation of τ a sequential experiment design is presented: the fluorescence model of Chapter 4 is used to design an information rich PAM experiment. The theoretical results considering *in silico* experiments are presented and discussed. Next, the practical identifiability with the real experiments is tested. Moreover, a constant actinic light experiment with a recovery phase one hour long will be used to estimate the recovery rate k_r. The model structure needs to be enhanced to represent the new data, leading to a more complex model. A second experiment design will be performed for the identification of the complex model. The effectiveness of the design will be tested from a theoretical and practical point of view. Although some experimental issues still need to be tackled, the results prove that experiment design can be very useful to increase the information contained in a PAM experiment and a confident estimation of 14 out of 16 parameters are estimated with good statistical significance. Some final remarks will conclude the Chapter.

6.2 Antenna size measurements

As discussed in Chapter 4 the parameter τ need to be fixed to a literature value, as PAM experiments are not suitable to investigate the very fast dynamic related to photoproduction. Here we want to introduce a different type of measurement, that can be coupled to the standard PAM experiments to gather the necessary information to estimate τ.

The fluorescence kinetics are measured with a LED pump and probe JTS10 spectrophotometer in the fluorescence mode. Fluorescence inductions are measured in the infra-red region of the spectrum upon excitation with blue light at 450 nm. 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) was added at a concentration of 80 μM to prevent oxidation of the primary quinone acceptor Q_A, i.e. the reaction rate from B to A in the Han model (Joliot and Joliot, 2006; Simionato et al., 2013b). In the presence of this inhibitor, the half saturation time of the fluorescence rise is inversely proportional to the so called ASII (Bonente et al., 2012). In the Han model ASII is equal to the product $\sigma_{PS2}I$ as the reaction from open to closed RCs is assumed to be a first order reaction with respect to the light intensity. The saturation curves are measured for different actinic light and are fitted with exponential curves to estimate the values of ASII. The analysis of these measurement will

be used both to verify the correctness of Han model hypothesis of first order reaction and to estimate parameter σ_{PS2} of the Han model.



(a) ASII experiment (b) ASII vs I

Figure 6.1: *(a) Fluorescence emission kinetics of PSII from dark-adapted acclimated cells were treated with DCMU. The time required for reaching half of the maximum is inversely proportional to the so-called functional PSII antenna size. The fluorescence flux are reported for three different actinic light and normalised by the maximum value (a.u.: arbitrary units). (b) The value of ASII is reported for five different light intensities along with the linear fit. The orange envelope represents the confidence region of the linear fitting.*

In Figure 6.1(a) an example of saturating curves for three different actinic light intensities has been reported. The experiment we considered has been performed for five different actinic light intensities (45, 80, 150, 320, 950 and 2050 μE/m^{2}s) and four repetition for each light intensities have been executed. The values of ASII for the different light intensities are reported in Figure 6.1(b). The variances are very small with a maximum value of approximately 2.5% of the ASII value at 2050 μE/m^{2}s. The slope of the fitting line is equal to 0.25 $\pm$ 0.009 with an R^2 value of 0.993. The orange envelop represent the confidence region of the linear fitting.

The fitting results reported in Figure 6.1(b) validate the Han assumption and provide useful information to enhance the parameter estimation of the fluorescence model: an additional experiment where the value of σ_{PS2} is a measurement with a constant variance equal to 0.009 has been added to the calibration set. Note that σ_{PS2} according to the fluorescence model is variable with NPQ activity, the measured value correspond to the value of σ_{PS2} for a dark adapted sample as the fast time horizon of the experiment, 0.3 seconds, is sufficiently small to prevent the activation of photoregulation mechanisms. With respect to the parameter estimation results of Chapter 4 the additional experiment allows to obtain a more accurate estimation of σ and S_F, moreover parameter τ can be estimated with good accuracy (see Table 6.1).

Table 6.1: *Effect of the introduction of σ_{PS2} measurements. The parameters not reported here are not significantly affected by the additional experiment. Normalised values of parameters are reported along with 95% confidence interval and t-values. The reference t-value is 1.65.*

Parameter	Estimate	95% conf. int.	t-value 95%
S_F	$2.06 \times 10^{\ 0}$	1.13×10^{-1}	18.32
σ	6.60×10^{-1}	3.51×10^{-2}	18.74
τ	7.28×10^{-3}	5.36×10^{-4}	13.58

6.3 Model based design of experiment to improve the parameter estimation of fluorescence model

Two new PAM experiments will be introduced in the calibration set to improve the accuracy of parameter estimates. In order to have a more precise estimation of parameters estimated in Chapter 4 a MBDoE approach will be used, next a constant light experiment with a recovery phase one hour long will be used to estimate k_r as additional parameter.

6.3.1 Optimal experimental design

The experiment considered for model calibration refers to a sample of N. Gaditana acclimated at 100 μE/m^{2}s(the experiment will be referred as Ch4Exp in the following). As discussed in Chapter 4 some of the parameters of the model have been fixed to a literature value, as they require specific experiments to be accurately estimated. Table 6.2 summarises the estimates of the parameters values as obtained by a parameter estimation in the case a standard (non-designed) PAM experiment is used. The confidence intervals and t-values are also reported in Table 6.2. Although, there is a very good agreement between experimental data and the model as presented in Chapter 4, from Table 6.2 it can be observed that the model is not accurately identified. In fact, the parameter estimation is not satisfactory from a statistical point of view, as some parameters are characterised by large confidence intervals (and low t-values). In particular, parameters k_d and η_I have a t-value well below the reference t-value, thus suggesting a correlation between the two parameters. However, C dynamics as represented by the third equation in set 4.3, shows the importance of parameter k_d at representing kinetics leading to damage in the reaction centres.Parameter k_d is therefore quite a significant parameter and its precise estimation is advocated for an accurate description of light induced inhibition.

Thus, in order to improve the precision of parameter estimation and to identify the model an MBDoE has been performed. The design is based on the A-criterion[2] and aims at optimising a PAM protocol with 20 light steps and 57 measurements. The optimisation determines the light intensity of each light step and the measuring points (i.e. the time at which a saturating

[2]Also D-criterion and E-criterion has been considered but they lead to numerical issues and for this reason the A-criterion has been preferred.

Table 6.2: *The first three column report parameter values estimated using one standard (non-designed) PAM experiment along with 95% confidence interval and t-values. Reference t-value is 1.65. The last three columns report the parameter estimates, confidence intervals and t-values using an optimally designed experiment. Statistically unsatisfactory estimates are indicated by (*).*

Par	Par. Value	95% c.i..	t-val 95%	Par. Value	95% c.i.	t-val 95%
	Non designed experiment			Optimally designed experiment		
ξ	5.83×10^{-2}	1.39×10^{-2}	4.19	5.73×10^{-2}	3.81×10^{-3}	15.05
I_{qE}	8.32×10^{2}	8.21×10^{1}	10.13	8.56×10^{2}	5.26×10^{1}	16.28
k_d	7.17×10^{-7}	1.08×10^{-6}	**0.67***	6.91×10^{-7}	2.35×10^{-7}	2.94
n	2.35×10^{0}	2.03×10^{-1}	11.59	2.34×10^{0}	1.48×10^{-1}	15.88
η_I	7.57×10^{1}	1.10×10^{2}	**0.69***	78.1×10^{1}	2.17×10^{1}	3.60
η_{qE}	1.91×10^{1}	1.93×10^{0}	9.95	1.96×10^{1}	1.00×10^{0}	19.58
η_P	1.13×10^{1}	3.53×10^{-1}	32.01	1.13×10^{1}	2.84×10^{-1}	39.91
S_F	1.65×10^{0}	2.49×10^{-1}	6.61	1.60×10^{0}	1.59×10^{-1}	10.09
σ	8.23×10^{-1}	1.25×10^{-1}	6.60	8.47×10^{-1}	8.40×10^{-2}	10.09

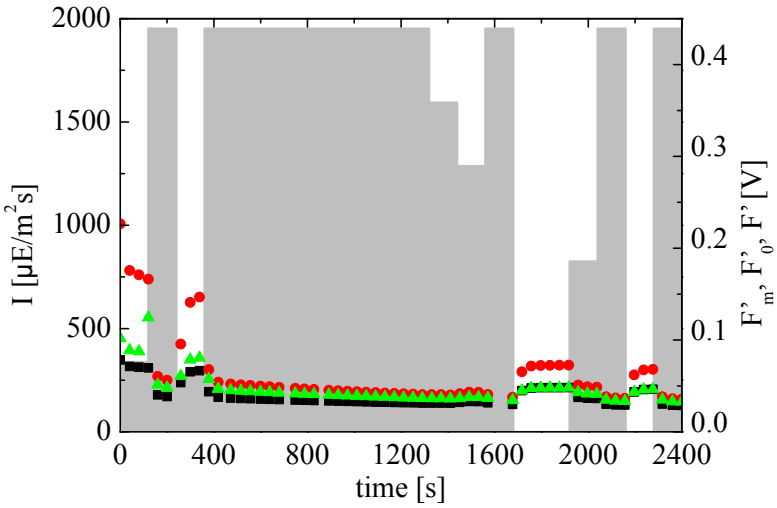
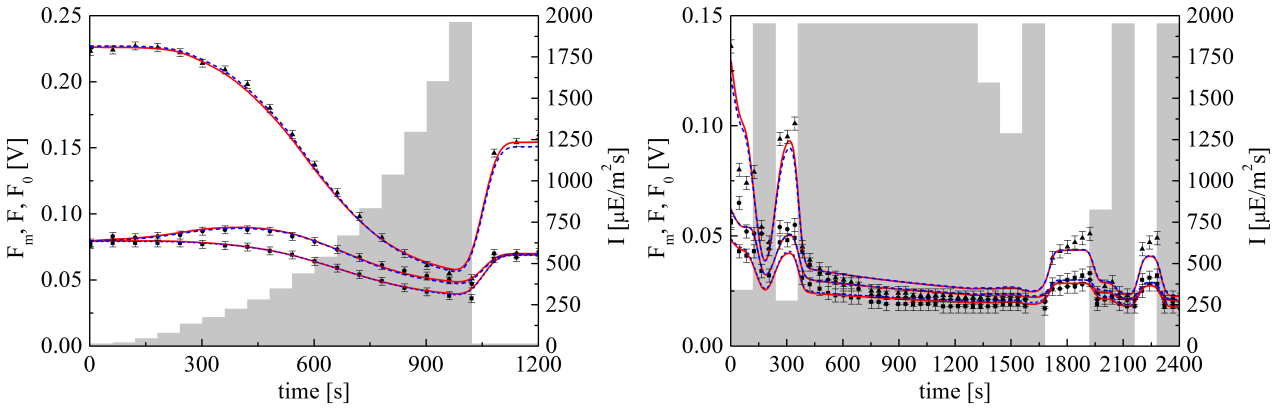


Figure 6.2: *Optimal designed experiment (DOE1) for the simple NPQ model for the two alternative experiment structure considered.*

pulse is applied). The minimum time gap between two measurement has been set to 40 s in order to assure the validity of the biological assumption that the saturating pulses do not affect photoinhibition and photoregulation. Finally, the actinic light profiles were approximated using the intensities that the PAM fluorometer can apply. In fact the AL in the PAM can assume only a particular set of values, i.e.: 0, 6, 13, 22, 37, 53, 70, 95, 126, 166, 216, 273, 339, 430, 531, 660, 825, 1028, 1287, 1594 and 1952 μE/m^{2}s. In Figure 6.2 the optimal experiment is reported along with the simulated measurements.

The last three columns of Table 6.2 report the newly estimated values of the parameters, the confidence intervals, and t-values after the designed experiment. The results show that a confident parameter estimation can be achieved through the utilisation of MBDoE. It is important to underline that measurement noise or model mismatch can hinder the practical



(a) Ch4Exp (b) DOE1

Figure 6.3: *Calibration results for the simple NPQ model and two different values of initial damage C_0: continuous lines consider $C_0 = 0$, dashed lines $C_0 = 0.01$*

identifiability. In the next section the experimental validation of the suggested procedure will be reported.

6.3.2 Practical model identifiability

Although the theoretical results seem to be promising, the practical identifiability with the real experiments has to be tested. The optimally designed experiment reported in previous section has been realised using a sample of Nannochloropsis Salina acclimated to 100 μE/m^{2}sand added to the calibration set. In Figure 6.3 the calibration results are reported for two different initial condition of damage: the continuous curves consider $C_0 = 0$ while the dashed lines consider $C_0 = 0.01$.

The fitting in the two scenarios is equivalent but if an initial value of damage is considered a better parameter estimation is achieved. The initial damage has been assumed because the value of F_v/F_m at the beginning of the protocol of the designed experiment was smaller than the value for Ch4Exp (0.58 instead of 0.625) indicating an initial inhibition of the sample. The parameter estimation results are reported in Table 6.3. The introduction of optimally design experiments lead to a better parameter estimation in terms of statistical significance. However, the mismatch between the model predictions and the experimental data lead to a significant decrease in parameter estimation accuracy. In fact, the model is in fairly good agreement with experimental data but tends to overestimate the fluorescence fluxes between 700 and 1500 seconds and fails to predict accurately the NPQ relaxation during the two dark phases at the end of the protocol. These mismatches lead to a great reduction of accuracy of parameter estimation with respect to the theoretical one. In particular, only if we assume an initial damage we can have a statistically meaningful estimation of the critical parameters k_d and η_I. These issues will be tackled in the next sections. However, first a procedure will be proposed for estimating k_r.

Table 6.3: *Normalised values of parameters are reported along with 95% confidence interval and t-values. The reference t-value is 1.65. The normalisation factors are the estimates in the first column of Table 6.2. Statistically unsatisfactory estimates are indicated by (*).*

Parameter	Normalised value		95% conf. int.		t-value 95%	
	$C_0 = 0$	$C_0 = 0.01$	$C_0 = 0$	$C_0 = 0.01$	$C_0 = 0$	$C_0 = 0.01$
ξ	2.05	2.03	2.71×10^{-1}	2.85×10^{-1}	7.56	7.118
I_{qE}	0.91	0.93	6.21×10^{-2}	6.65×10^{-2}	14.65	14.01
k_d	0.50	0.94	3.95×10^{-1}	4.02×10^{-1}	**1.255***	2.351
n	1.00	1.02	7.45×10^{-2}	7.50×10^{-2}	13.37	13.61
η_I	1.97	1.21	1.47×10^{0}	4.87×10^{-1}	**1.337***	2.49
η_{qE}	0.95	0.98	5.79×10^{-2}	6.66×10^{-2}	16.41	14.72
η_P	0.98	1.00	2.63×10^{-2}	2.59×10^{-2}	37.06	38.5
${S_{F1}}^a$	0.93	0.98	1.07×10^{-1}	1.18×10^{-1}	8.63	8.275
${S_{F2}}^a$	0.58	0.64	6.91×10^{-2}	8.45×10^{-2}	8.36	7.601
σ	1.09	1.04	1.26×10^{-1}	1.25×10^{-1}	8.63	8.287

[a] As the two PAM experiments have very different absolute values of fluorescence flux due to different cells concentration in the samples a scaling factor for each experiment has been used. S_{F2} refers to the optimally designed experiment.

6.3.3 Estimating recovery of inhibited PSUs

We want to tackle the problem of estimating k_r using a standard (non designed) experiment. In particular, the experiment used to estimate k_r (that in the following will be indicated as RecExp) is an experiment where a strong actinic light of 2000 μE/m^{2}shas been kept constant for 10 minutes in order to induce a certain amount of damage in the photosynthetic apparatus, after that a dark phase of one hour has been considered to analyse the recovery, thus estimating k_r. Before running the experiment, a simulated experiment has been considered to assure that one hour of dark phase is enough to obtain a statistically meaningful estimation of k_r. However, the realised experiment underlined a strong model mismatch that hinder the possibility of estimating k_r and will be analysed in detail in the following section where the fluorescence model will be enhanced to better describe the dynamic of the recovery phase. In Figure 6.4(a) the experiment RecExp is compared with the predicted fluorescence fluxes.

6.4 Model enhancing: two steps NPQ

The experiment in Figure 6.4(a) is necessary to estimate accurately the parameter k_r that represent the recovery rate of inhibited PSUs. However the fluorescence model developed so far is not able of representing those data. In fact, the recovery phase suggests a bi-exponential behaviour of NPQ relaxation as we can observe in Figure 6.4(b), while the current model represent it with a simple exponential curve where ξ determines the time constant of the relaxation process. A new process with time scales of minutes needs introducing. Note that the recovery of inhibited PSUs must act on a time scales of hours and a fast (time scale of

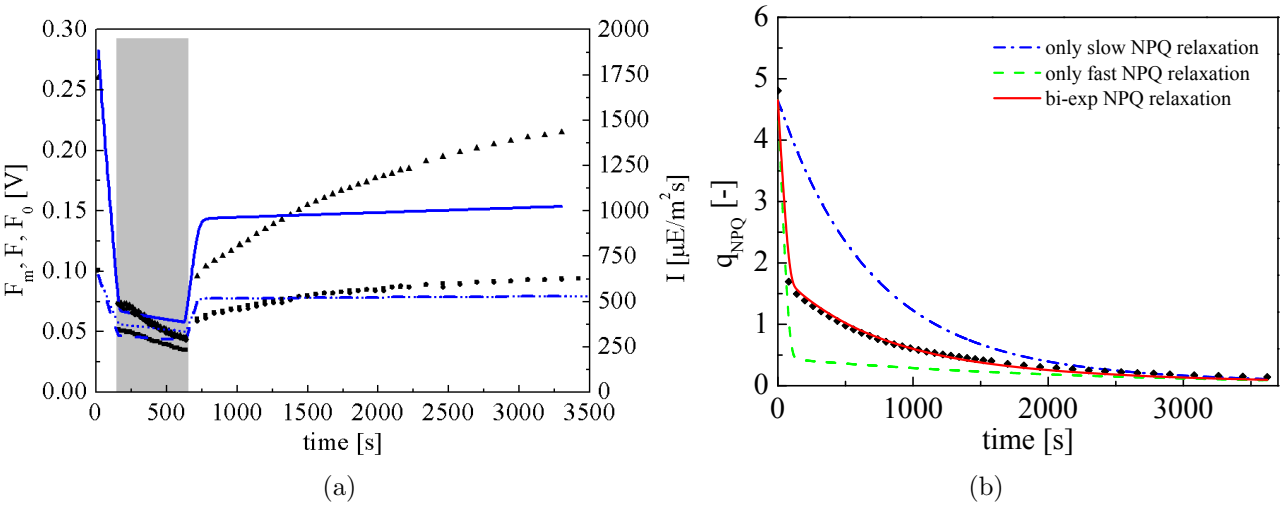


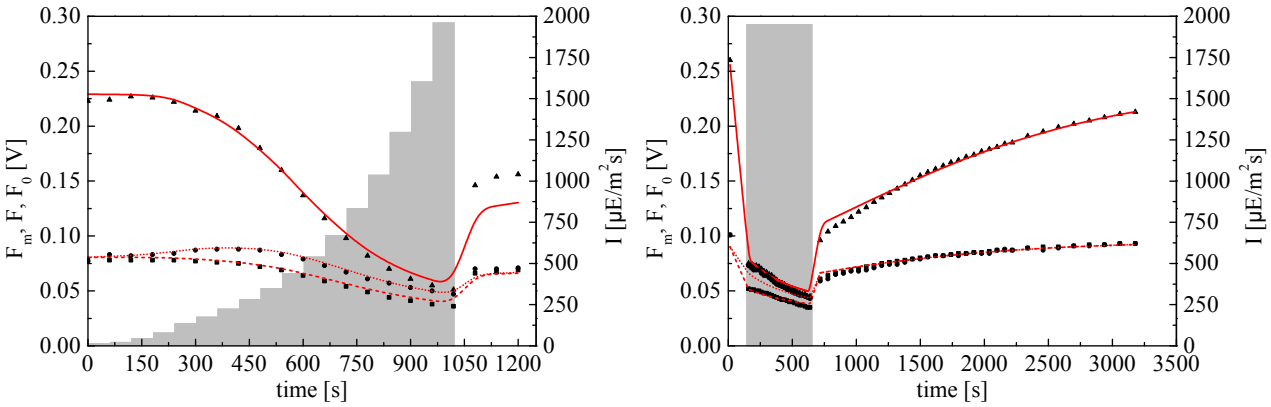
Figure 6.4: *(a) Comparison between the predicted and measured fluorescence fluxes F_m' (triangles), F_0' (squares) and F' (circles) in response to a constant light experiment. The grey-shaded area represents the light intensity. (b) Measured value of q_{NPQ} during the recovery phase of experiment RecExp along with predicted values using a single exponential curve or a biexponential curve.*

seconds) component of NPQ is required to explain the recovery in the first minute of the dark phase (the interval between the last measurement point of the light phase and the first one of the dark phase).

From a biological point of view, this is explained by the fact that the NPQ is related to two main processes. The first one is a fast process that involves the activation of the LHCSR protein. This process has a time constant of seconds for both activation and relaxation. The second one is related to the zeaxanthin and acts in time scales of minutes. Moreover, zeaxanthin has a complex effect on the NPQ activity: it both enhances the quenching effect of LHCSR and acts as an additional quencher (Pinnola et al., 2013). Accordingly, we can express NPQ activity as:

$$\eta_{qE} = \alpha_F \left(\overline{\eta}^F_{\mathrm{qE}} + \alpha_S \, \overline{\eta}^{int}_{\mathrm{qE}}\right) + \alpha_S \, \overline{\eta}^S_{\mathrm{qE}} \tag{6.1}$$

As both LHCSR and zeaxanthin activity are activated by low lumen pH, we can make the hypothesis that the reference activity function of α_F and α_S is the same and equal to α_{SS}. Using this hypothesis we do not need to have two independent Hill functions but the complexity of the model is still high as we introduce three additional parameters: the time constants of the slow NPQ process ($\overline{\eta}^S_{\mathrm{qE}}$), and the parameters $\overline{\eta}^{int}_{\mathrm{qE}}$ and $\overline{\eta}^S_{\mathrm{qE}}$ that represent the quenching effects of the slow NPQ process. Moreover, the estimation problem is more difficult as the three recovery processes partially overlap, while in the original model NPQ relaxation and damaged PSUs recovery were acting on very different time scales. A calibration of the new model has been attempted considering the Ch4Exp along with RecExp. Moreover, the ASII experiment has been included in order to estimate parameter τ. Figure 6.5 the calibration results are reported showing a good agreement between the experimental data and the model. In Table 6.4 are reported the parameter values along with 95% confidence intervals and t-values.



(a) Ch4exp (b) RecExp

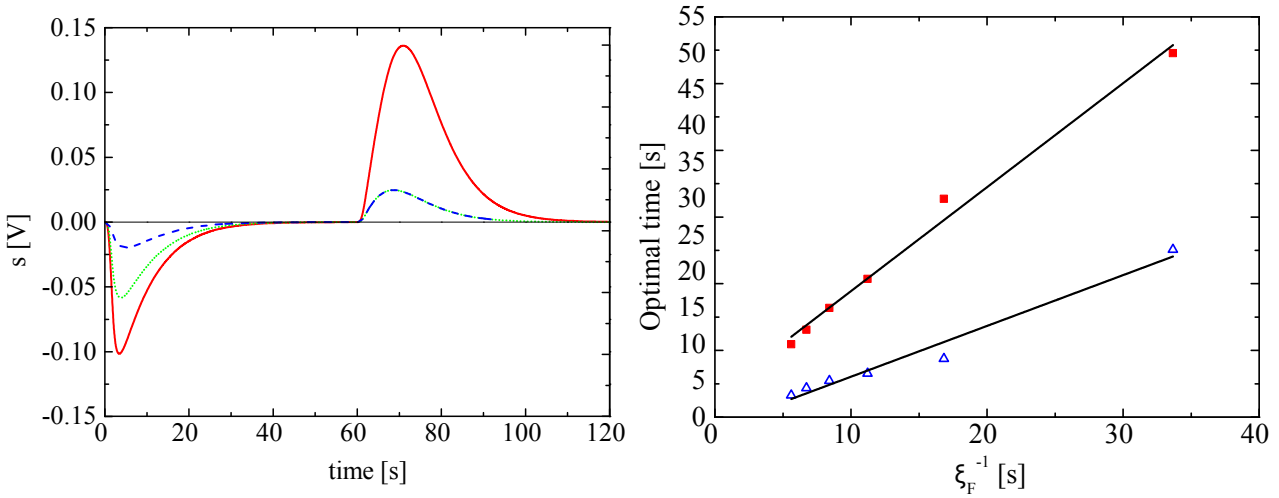
Figure 6.5: *Calibration results for complex NPQ model. Measured fluorescence fluxes F'_m (triangles), F'_0 (squares) and F' (circles) are reported along with model predictions. The grey shaded area represents the light intensity.*

Table 6.4: *Parameter estimates are reported along with 95% confidence interval and t-values. The reference t-value is 1.65. The calibration set is composed by Ch4Exp, RecExp and σ_{PS2} measurement.*

Parameter	Estimated value	95% conf. int.	t-value 95%
ξ_F	1.86×10^{-1}	1.95×10^{1}	**0.0095 ***
ξ_S	9.74×10^{-4}	6.71×10^{-5}	14.53
I_{qE}	5.99×10^{2}	4.20×10^{1}	14.26
k_d	2.16×10^{-6}	1.30×10^{-6}	**1.62 ***
n	2.18×10^{0}	1.75×10^{-1}	12.56
η_I	3.05×10^{0}	1.56×10^{0}	1.954
$\overline{\eta}^F_{\text{qE}}$	8.20×10^{0}	9.21×10^{-1}	8.9
$\overline{\eta}^S_{\text{qE}}$	1.91×10^{1}	1.19×10^{0}	16.1
$\overline{\eta}^{int}_{\text{qE}}$	2.43×10^{1}	3.32×10^{0}	7.31
η_P	1.13×10^{1}	3.06×10^{-1}	37.02
$S_{F1}{}^a$	2.09×10^{0}	1.17×10^{-1}	17.83
$S_{F2}{}^a$	2.32×10^{0}	1.38×10^{-1}	16.87
σ	6.61×10^{-1}	3.50×10^{-2}	18.85
τ	8.35×10^{-3}	1.07×10^{-3}	7.808

[a] S_{F1} refers to Ch4exp, S_{F2} refers to RecExp experiment. The different values are due to different cells concentration of the samples.

With the new expression of NPQ dynamic the model is able to reproduce in a satisfactory way the experimental data. However due to the higher complexity of the model a number of identifiability issues arise: (i) parameters ξ_F and k_d fail the t-test and in particular ξ_F has huge confidence intervals and is not identified at all; (ii) parameter η_I has a t-value greater than the reference t-value but the confidence intervals are still quite large; and, (iii) parameter k_r can not be estimated and for this reason has been kept fixed. The unidentifiability of parameter k_r, despite the utilisation of RecExp, is due to the fact that the dynamics of the slow NPQ

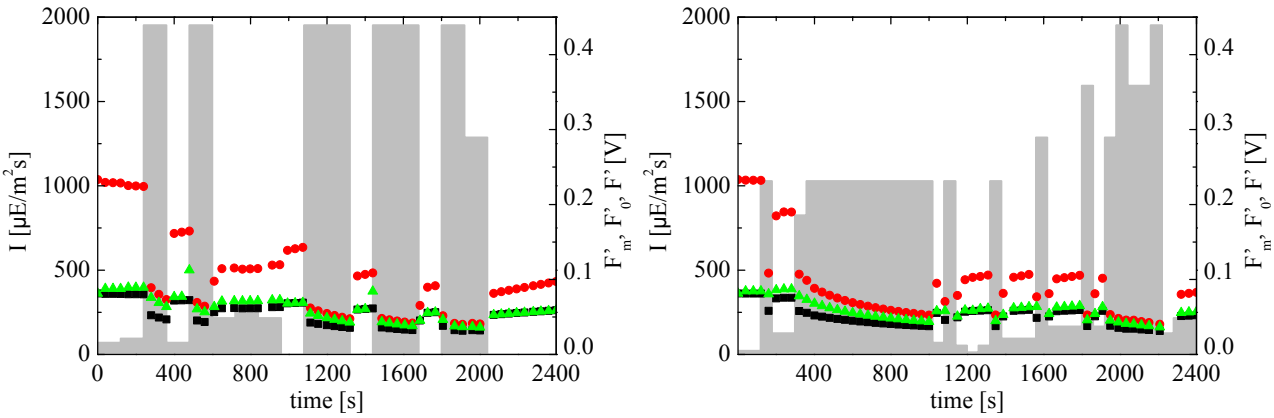


(a) (b)

Figure 6.6: *(a)Sensitivity profile of parameter ξ_F. The red continuous line represent the sensitivity for F'_m, the green dotted line represent the sensitivity for F' and the blue dashed line represent the sensitivity for F'_0. The protocol to obtain those curves was 60 seconds of strong AL followed by 60 seconds of dark. The actual value of ξ_F used was 0.18 s^{-1}. (b) Dependence of optimal time for the sensitivity of F'_m with respect to the reverse of ξ_F The square represent the optimal sampling time after that the light is switched off, the triangles represent the optimal sampling time after that the light is switched to 2000 μE/m^{2}s.*

partially overlap with the recovery of inhibited PSU. The overlapping of the two relaxation processes requires a longer dark phase to discriminate between the two processes.

In order to better identify the model, an MBDoE approach will be used. The parameter ξ_F, is particularly difficult to identify due to the rapid dynamics of the fast NPQ processes. An important issue to underline is the practical identifiability of parameter ξ_F. In Figure 6.6(a) we report the sensitivity profile for a case study where the light is switched from zero to 2000 μE/m^{2}sand after 60 seconds is switched off. The sensitivity is calculated as $(y' - y)/\Delta\theta$ where y and y' are the measured variables calculated with the nominal value of ξ_F and with a perturbed one respectively, while $\Delta\theta$ is the parameter perturbation. If we look at the profiles we can observe that: (i) the most sensitive variable is F'_m followed by F', (ii) the maximum (or minimum) sensitivities are reached after few seconds. Moreover, the time when the maximum is reached strongly depends on the value of ξ_F as showed in Figure 6.6(b). The strategy is to tune the measuring points after a big step variation of actinic light in order to improve the accuracy of estimation of parameter ξ_F without affecting the accuracy of the estimation of the other parameters. Having some measurements after 10 seconds after a step variation from high irradiance to dark has shown to be a robust strategy for a wide range of ξ_F that corresponds to a time scale from approximately 5 to 30 seconds.



(a) DOE2 (b) DOE3

Figure 6.7: *Optimal designed experiments for complex NPQ model.*

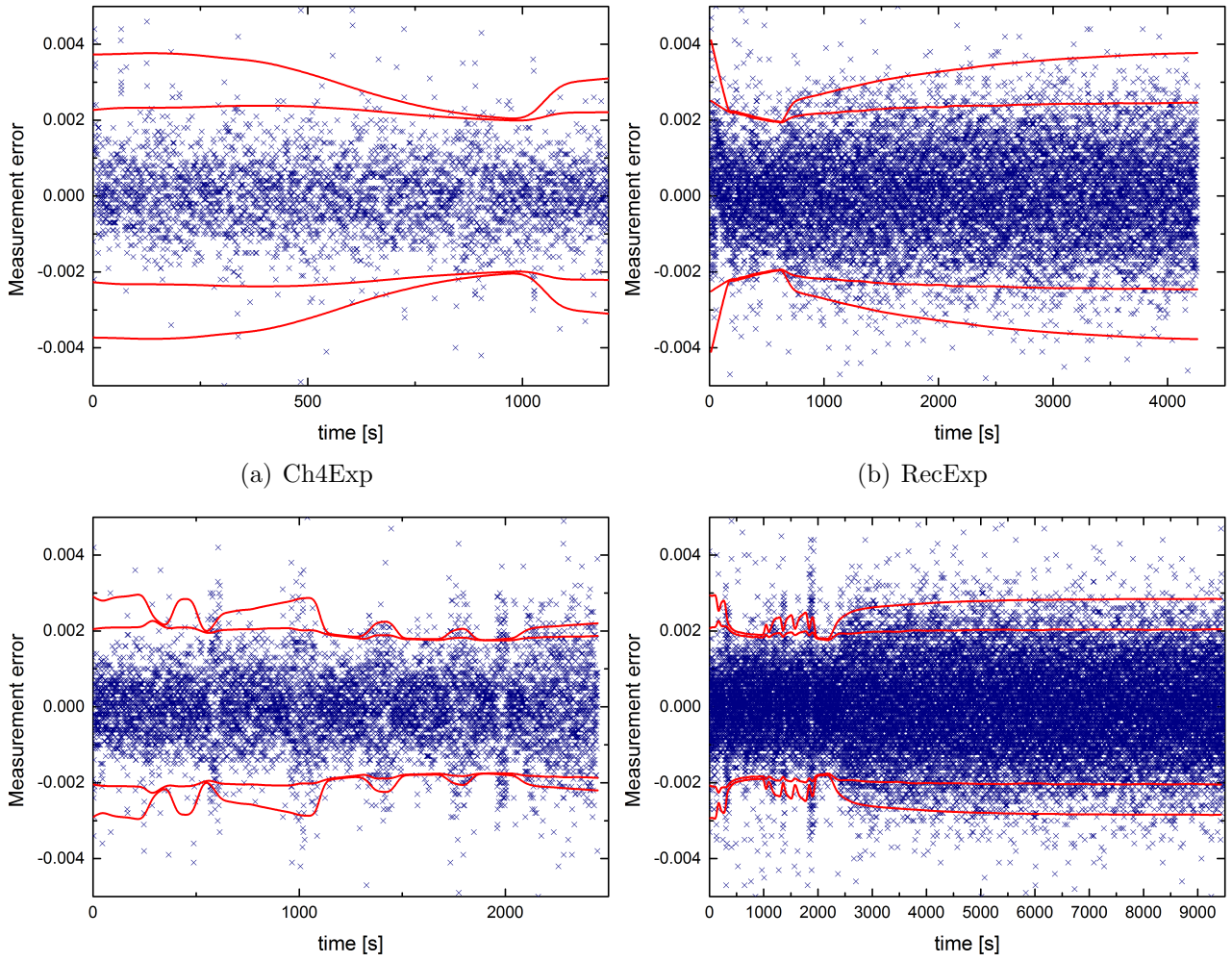
6.5 Design of experiment for the complex NPQ model

The design considers two experiments 2400 seconds long with 57 measurement points, optimised under the constraint of 40 seconds minimum between two measurements. The first experiment has 20 equidistant control intervals while the second has 23 intervals of variable length. In Figure 6.7 are reported the optimal AL profiles and the predicted experimental points. In Table 6.5 are reported the new estimates (normalised by the values in Table 6.4) considering the optimally designed experiments along with 95% confidence intervals and t-values.

Table 6.5: *Normalised parameter values, 95% confidence interval and t-values for a calibration set composed by Ch4exp, RecExp one of the optimally designed experiments and σ_{PS2} measurement. The reference t-value is 1.65.*

Par.	Norm. val.	95% conf.int.	t-val 95%	Norm. val.	95% conf.int.	t-val 95%
	DOE2 optimal exp			DOE3 optimal exp		
ξ_F	0.85	8.82×10^{-2}	9.60	0.88	6.44×10^{-2}	13.64
ξ_S	1.02	5.47×10^{-2}	18.65	0.98	5.23×10^{-2}	18.71
I_{qE}	1.10	5.99×10^{-2}	18.38	1.08	6.13×10^{-2}	17.69
k_d	0.70	1.04×10^{-1}	6.69	0.85	1.21×10^{-1}	7.02
n	0.89	5.15×10^{-2}	17.35	0.88	4.03×10^{-2}	21.77
η_I	2.52	5.61×10^{-1}	4.49	1.35	5.56×10^{-1}	2.42
$\overline{\eta}^F_{\mathrm{qE}}$	1.07	6.67×10^{-2}	16.09	1.02	7.60×10^{-2}	13.42
$\overline{\eta}^S_{\mathrm{qE}}$	1.10	5.91×10^{-2}	18.58	1.09	5.07×10^{-2}	21.51
$\overline{\eta}^{int}_{\mathrm{qE}}$	1.18	1.10×10^{-1}	10.77	1.19	9.33×10^{-2}	12.77
η_P	1.00	2.02×10^{-2}	49.93	1.00	2.00×10^{-2}	50.04
${S_{F1}}^a$	0.98	5.17×10^{-2}	18.99	0.99	1.24×10^{-2}	80.35
${S_{F2}}^a$	1.14	6.05×10^{-2}	18.90	1.12	2.02×10^{-2}	55.49
σ	1.03	5.23×10^{-2}	19.65	1.01	1.45×10^{-2}	69.89
τ	1.04	1.08×10^{-1}	9.57	0.98	1.05×10^{-1}	9.31

[a] S_{F1} refers to Ch4exp, DOE2 and DOE3, S_{F2} refers to RecExp experiment. The different values are due to different cells concentration of the samples.



(a) Ch4Exp (b) RecExp

(c) DOE2 (d) DOE3

Figure 6.8: *Experimental error in long recovery experiment 6.8(b), Ch4Exp 6.8(a), DOE2 6.8(c) and DOE3 6.8(d) along with error profiles predicted by the linear variance model considered in this chapter. The two profiles refers to $F_m^{'}$ and $F^{'}$, each profile is simmetric with respect to the x-axis.*

We can observe that both experiments lead to very accurate parameter estimations, at least from a theoretical point of view. In experiment DOE3 a dark period of two hours has been added to gather some additional information for the estimation of k_r. The realised experiments will be used in the next section to validate the theoretical results.

6.5.1 Measurement error evaluation

The fluorescence flux is measured by the PAM every 0.3 seconds and the variation between two consecutive points, if we exclude the saturating pulses, can be considered an estimation of the measurement noise. In Figure 6.8 is reported a scatter plot for each experiment with the measurement noise as a function of time. Moreover, the error bars calculated by the variance model utilised in this chapter are reported for $F_m^{'}$ and $F^{'}$. The profiles representing the error bars are symmetrical and the value for $F_m^{'}$ always has a bigger absolute value, as $F_m^{'}$ is always

Table 6.6: *Effect of MBDoE: Parameter estimates are reported along with the normalised values (with respect to the values reported in Table 6.4), 95% confidence interval and t-values. The reference t-value is 1.65. Calibration set A includes the experiments Ch4Exp and DOE3; calibration set B includes the experiments Ch4Exp, RecExp and DOE3.*

Par.	Norm. val.	95% conf.int.	t-val 95%	Norm. val.	95% conf.int.	t-val 95%
	Calibration set A			Calibration set B		
ξ_F	1.20	1.08×10^{-1}	11.12	1.44	1.89×10^{-1}	7.66
ξ_S	1.39	1.01×10^{-1}	13.75	1.36	7.30×10^{-2}	18.61
I_{qE}	0.80	5.21×10^{-2}	15.29	0.99	3.43×10^{-2}	28.84
k_d	0.41	1.62×10^{-1}	2.55	0.43	1.24×10^{-1}	3.46
k_r	1.00	-	-	1.01	5.28×10^{-1}	1.91
n	1.11	9.12×10^{-2}	12.13	1.11	5.89×10^{-2}	18.84
η_I	4.56	1.85×10^{0}	2.47	5.00	1.52×10^{0}	3.29
$\overline{\eta}^F_{\mathrm{qE}}$	0.41	6.20×10^{-2}	6.58	0.72	6.06×10^{-2}	11.91
$\overline{\eta}^S_{\mathrm{qE}}$	0.38	1.95×10^{-2}	19.48	0.64	3.03×10^{-2}	20.97
$\overline{\eta}^{int}_{\mathrm{qE}}$	1.24	8.09×10^{-2}	15.31	1.01	6.88×10^{-2}	14.62
η_P	0.91	2.20×10^{-2}	41.33	0.91	2.06×10^{-2}	44.27
${S_{F1}}^a$	1.93	1.08×10^{-1}	17.90	1.92	1.05×10^{-1}	18.18
${S_{F2}}^a$	1.34	7.52×10^{-2}	17.87	2.19	1.21×10^{-1}	18.18
${S_{F3}}^a$	-	-	-	1.38	7.58×10^{-2}	18.28
σ	1.03	5.48×10^{-2}	18.74	1.04	5.46×10^{-2}	19.01
τ	0.70	8.72×10^{-2}	7.99	0.83	8.99×10^{-2}	9.24

[a] S_{F1} refers to Ch4exp; S_{F2} refers to RecExp; S_{F3} refers to DOE3 experiment. The different values are due to different cells concentration of the samples.

bigger than $F^{'}$ by definition.

We can observe that the variance model represents the measurement error for the four experiments in a reasonable way. On the contrary the variance model used in Chapter 4 that can be obtained subtracting the constant term from the reported profiles clearly underestimates the measurement errors. Another alternative is to use a constant variance model; however, it makes sense that the measuring noise has a component proportional to the signal intensity.

6.5.2 Practical model identifiability

In order to identify the model, the designed experiments have been performed and added to the calibration set. Experiment DOE2 leads to parameter values very different from the values reported in Table 6.4 and in particular activity of NPQ seems to be quite peculiar. The reason of the mismatch may be related to some experimental issues and for this reason the experiment will not be used for calibration. Future experiments will replicate experiment DOE2 in order to understand the reason of the mismatch. In the following, only experiment DOE3 is considered as additional calibration experiment.

In order to have a fair comparison between the standard and the optimally designed PAM experiments we have compared a calibration set composed by two standard PAM experiments,

Ch4exp and RecExp (Table 6.4) with a calibration set composed by the experiment of Ch4Exp and the optimally designed experiment DOE3. Both calibration sets include the ASII experiment to estimate τ. The parameter estimation results for the optimal calibration set are reported in Table 6.6 and the predicted fluorescence profiles for Ch4Exp and DOE3 are reported in Figure 6.9(a) and 6.9(c).

The results in Figure 6.9 show an excellent agreement between experimental data and model predictions. Moreover, the parameter are estimated with good accuracy. From Table 6.6 we can observe how the parameter ξ_F is now estimated in a reliable way, with a confidence interval 95% smaller than 10% of the estimated value. The parameters k_d and η_I are also estimated with higher accuracy, even if from the theoretical results a better estimation of k_d was expected.

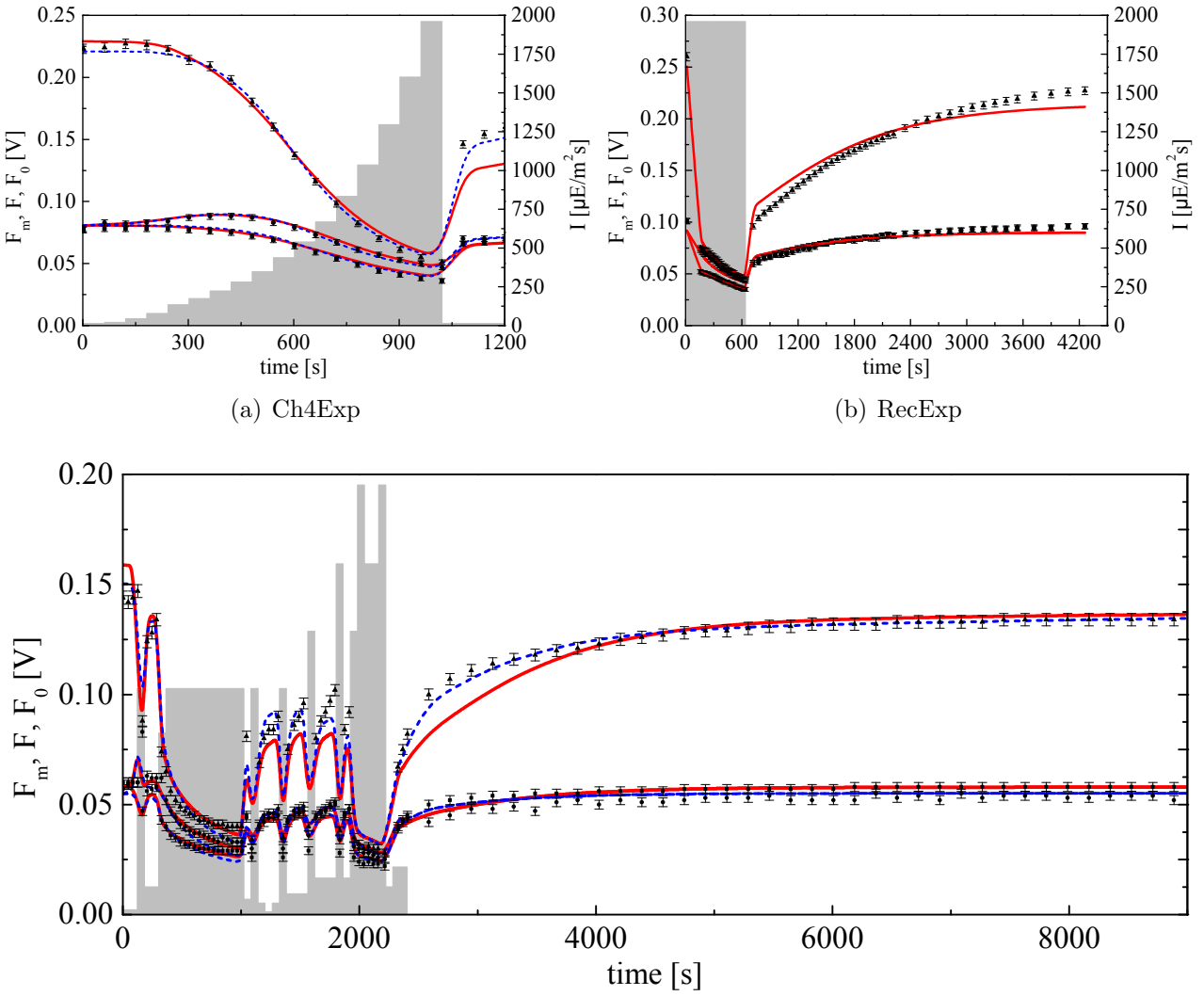
6.5.3 Estimating recovery of inhibited PSUs

At this point, it is eventually possible to proceed with the estimation of parameter k_r. With respect to the simple NPQ model here the problem is much more complex as the slow NPQ process partially overlap the damaged PSUs recovery process. In order to maximise the information about recovery processes the two experiments with long recovery phase (i.e. RecExp and DOE3) need considering. A calibration set composed by Ch4Exp, RecExp, DOE3 and ASII experiment will be used. The parameter estimation results are reported in Table 6.6 and the model prediction are reported in Figure 6.9.

We can observe that the model represent with sufficient accuracy the data. The mismatch between experimental data and model prediction is mainly due to the fact that the RecExp experiment is slightly different from the other two (we can consider this difference due to biological variability). This point needs to be further analyse to determine which processes are more influenced by the biological variability and what are the expected range of variability. Although some minor problems are still present and related to the intrinsic variability that a biological system always has the model is able to reproduce experiments that are very different from each other and a statistically meaningful parameter estimation can be obtained. Parameter k_r is the less precisely identified, with a confidence interval approximately equal to 50% of the nominal value. To further improve the estimation of this parameter, even longer recovery phases need to be considered.

6.6 Complex NPQ model validation

The sequential MBDoE approach we followed has guided the model enhancement and led to a statistically meaningful parameter estimation. The experiments not used in calibration can now be used to validate the model. In order to test several light conditions the experiment used in Chapter 4 for calibration will be used, along with a constant light experiment. Moreover, experiments DOE1 and DOE2, not used in calibration will be used for validation purpose. The



(a) Ch4Exp

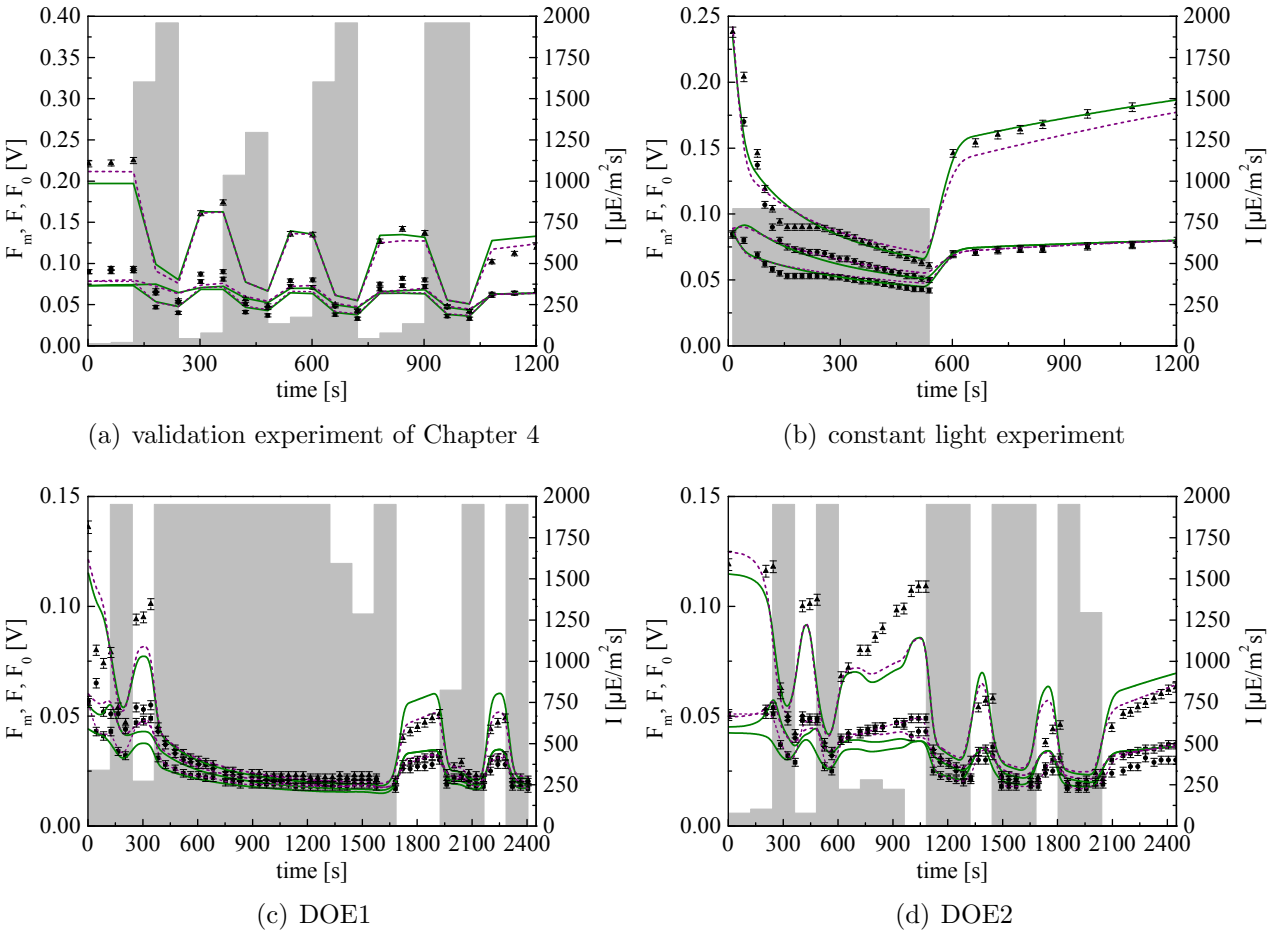
(b) RecExp

(c) DOE3

Figure 6.9: *Fluorescence profiles predicted by the model with two calibration experiments (dashed blue profile) and three calibration experiments (continuous red profiles).*

alternative model calibrations reported in Table 6.6 are both shown in Figure 6.10.

The continuous line refers to parameter values reported in Table 6.6 obtained with the calibration set A while the dotted line refers to the parameter values obtained with the calibration set B. The main difference between the two sets of predicted profiles can be observed in the dark periods of the validation experiments suggesting that parameter values obtained with the calibration set B are better than the values obtained with the calibration set A and thus underlying the reliability in the k_r estimation. The main model mismatch is present in the time interval between 600 and 1080 seconds of experiment DOE2 where a value of $F_m^{'}$ 20% smaller that the measured value is predicted. However, the results are overall very good showing that the model is able of accurate prediction for a wide range of PAM experiments. To our knowledge, there are no literature models that have been tested through such challenging experimental trials and have been proved to perform with such a consistent quality.



(a) validation experiment of Chapter 4

(b) constant light experiment

(c) DOE1

(d) DOE2

Figure 6.10: *Validation of florescence model. The continuous line refers to parameter values reported in Table 6.6 obtained with the calibration set A, the dotted line refers to the parameter values obtained with the calibration set B.*

6.7 Conclusion

In this chapter we proposed a model based design of experiments (MBDoE) approach to increase the accuracy of parameter estimation in fluorescence model. The measurement error has been described with a linear variance model and the assumption has been tested against experimental data. New measurements has been added to the calibration set to increase the number of estimable parameters. In particular, antenna size measurements has been used to investigate the fast dynamic process related to the photoproduction. On one hand, the antenna size measurements provided a clear verification of one of the hypothesis of the Han model, on the other hand the parameter τ related to the dynamic of photoproduction has been confidently estimated. Moreover, optimally designed experiments have been carried out to have a more precise parameter estimation. The new data required a more detailed representation of NPQ mechanism that reflects the actual biological process that is composed by two interdependent mechanisms. Experiments design provided information rich PAM experiments that have been proved to lead to accurate parameter estimation even if the complex NPQ representation is

considered.

The developed model has been validated against both constant light and variable light PAM experiments showing good agreement with experimental data. The results show that the NPQ dynamics can be investigated in an accurate way using PAM fluorometry. Longer time scales processes like damaged PSUs recovery can be studied as well but PAM experiments with a recovery phase longer than two hours are necessary to shrink the confidence intervals of the PSUs reparation rate.

Chapter 7

Conclusions and future perspectives

Microalgae are among the most promising alternative to substitute or at least to complement fossil fuels in the transport sector. However, the technical feasibility as well as the economics and environmental benefits of large-scale algal cultivation for biodiesel production are still matter of debate. Of particular concern, many economic and life cycle assessment are based on crude extrapolation of productivities obtained in the lab, where conditions differ drastically from those in outdoor culture systems, and no pilot or larger-scale demonstration plant has been able to reproduce them as of yet.

In this context, mathematical modelling can be a great help for developing a better understanding, and in turn enabling a better prediction capability, of microalgae culture dynamics. Models that convey state-of-the-art scientific knowledge are invaluable tools for unveiling and untangling the underlying photosynthetic and metabolic mechanisms. These models can be tested in a systematic way through dedicated experiments and, conversely, they can be used to guide the design of dedicated, information-rich experiments. For process development purposes too, models can be used to improve the design, operation and control of a microalgae culture system in order to enable and sustain a higher productivity.

7.1 Summary of thesis achievements

Although several modelling framework have been proposed in the literature, they still suffer from some limitations. In fact, many literature models are tailored to some experimental data, but their application out of the context in which they were developed is often infeasible. This is due to two main reasons: the first one is a lack of a general modelling development scheme that aims to assure the model identifiability; the second one, is the fact that many literature models are based on not fully verified biological assumptions and this can hinder their predictive capability. Hence, the first objective of this Thesis was to propose a general modelling development methodology, which can provide a useful tool to guide the model build-up. Moreover, we wanted to better exploit the information obtained by chlorophyll fluorescence measurements (in particular PAM fluorometry has been considered). Traditionally, a number

of fluorescence indexes have been used for monitoring specific photosynthetic mechanisms, by qualitatively relating these mechanisms to the measured fluorescence fluxes. In contrast, little effort has been devoted to quantifying these relations in the form of mathematical models, which would enable accurate predictions of the quantum yield of photosynthesis and in particular of its dynamic response to variable light conditions. In the following the main achievements of this Thesis are summarised.

(i) On the methodological side, a step-by-step approach has been presented and applied to guarantee the identifiability of the proposed models. The methodology has been applied to a case study where, some existing experimental data may be exploited to develop a model describing the growth as function of light intensity. A literature review has been carried out and two alternative models have been selected and discriminated to determine the best candidate model. The selected model has been enhanced to better represent some biological processes and a rigorous identifiability analysis has been performed on the modified model. The proposed methodology proved to be effective in selecting the best candidate model and to overcome the identifiability issues that arose. The results suggest that the developed model is accurate enough to represent all major processes of photosynthesis, photochemistry, PSU damage and energy dissipation. While results in reproducing experimental data are fully satisfactory, it should be underlined that algae growing in an industrial scale photobioreactor are exposed to different conditions. In particular light is not homogenously distributed because of cells shading, and illumination intensity is not constant because of diurnal changes and cells mixing. Finally nutrient availability can also be limiting and future efforts will be made to expand the model to include these phenomena, also by designing appropriate experiments. Moreover, although the identifiability analysis assures that the parameter set can be determined in a precise way the model is still based on some critical biological assumption. In particular the maximum quantum yield of photosynthesis is assumed to be linearly related to the fraction of inhibited PSUs. This hypothesis has already been used in literature but there is not a clear consensus regarding its validity. This motivates the following of the work, along with the necessity to represent some biological processes in a more general way, e.g. the energy dissipation via NPQ has only been accounted for in a static way.

(ii) A dynamic model of chlorophyll fluorescence has been developed and calibrated against PAM experimental data. The model accounts for photoproduction, photoregulation and photoinhibition in a semi-mechanistic way and aims at exploiting the quantitative informations that can be inferred from chlorophyll fluorescence. The model has been calibrated using a standard PAM experiment where the light follows an increasing step profile and proved to predict in an accurate way the fluorescence fluxes when validated against a challenging PAM protocol where high light periods are alternated with dark periods. The validation results suggest that the model grasp in a reliable way the main

processes acting on time scales from milliseconds to hours. The models capability to predict, in quantitative way, the state of the photosynthetic apparatus in terms of its open, closed and damaged reaction center has been assessed. This made it the first model of its kind capable of reliable predictions of the levels of photoinhibition and NPQ activity without the need for dedicated experiments (Ruban and Murchie, 2012), addressing a long-standing challenge in the modelling of photosynthetic productivity. Moreover, the availability of mathematical model that represent directly the fluorescence fluxes provide a tool to investigate the quantitative relationship between the fluorescence indexes and the state of photosynthetic apparatus. In particular, we have demonstrated that the linear relationship between q and the fraction of inhibited PSU, assumed by Wu and Merchuk (2001) and used also in Chapter 3, is valid only if the quenching effect of inhibited PSU is equal to the quenching effect of open reaction centres.

(iii) Two promising applications of the developed fluorescence model have been analysed in this Thesis. The first one, concerns the prediction of PI curves, based on chlorophyll fluorescence measurements and the possibility to set up a cross-validation framework, whereby both fluorescence and photosynthesis rate experiments could be used for model validation. The second one, is the utilisation of a MBDoE approach to design information rich PAM experiments. The first point has been analysed in Chapter 5. The fluorescence model has been extended to account for the photoacclimation process by varying a subset of the model parameters. The available data were not sufficient to propose robust empirical relationships to describe the variations of acclimation-dependent parameters. However, the results showed the PI curves predicted by our model are very similar to published literature data and a preliminary model validation has been carried out with experimental data from Gentile and Blanch (2001). Once validated the model allows to predict PI curves based on fast and reliable fluorescence measurements, thus avoiding the usual - and somewhat problematic - oxygen productivity measurements. The dynamic simulation of PI curves also underlines an important issue regarding the utilisation of the PI measurements. Typically in the literature PI curves are considered to be independent from the protocol used to obtain them; on the contrary, our results clearly show that the this assumption could lead to a significant overestimation of the productivity. The second issue has been addressed in Chapter 6. A MBDoE approach has been used to improve the accuracy of parameter estimation and for further testing the model structure. The new experiments revealed that the dynamic behaviour of NPQ cannot be described by a simple first order process, as initially assumed, but is the results of two interdependent mechanisms acting on different time scales. In fact, the two time scale NPQ mechanism is necessary to describe the dynamical transition between the dark adapted and the light adapted state if a constant light PAM experiments is considered. The optimally designed experiments allowed a statistically satisfactory parameter estimation and the final model

has been validated against several PAM protocols, showing a very good agreement with the experimental evidences. To our knowledge there are no literature models that have been tested through such challenging experimental trials and have been proved to perform with such a consistent quality.

(iv) One of the limitation of the fluorescence model was the necessity to fix to a literature value the parameter τ, which represents the time constant of the reaction rate from closed to open reaction centres, thus representing the dynamic of photoproduction. In fact, the time resolution of PAM measurement is too low to investigate the fast dynamic of photoproduction, which acts in a time scale of milliseconds. In order to confidently estimate τ antenna size measurements, obtained with a LED pump and probe JTS10 spectrophotometer in the fluorescence mode, have been considered. The utilisation of antenna size measurements provided an effective solution to achieve an accurate estimation of parameter τ. Moreover, one of the assumption on which is based the Han model and other state models is that the reaction rate from open to close reaction centres is a first order reaction with respect to the light intensity. The analysis of antenna size measurements at different light intensities allowed us to have a clear verification of this hypothesis.

7.2 Future perspective

Some considerations about future research challenges will be detailed in the following.

A first issue to address is to incorporate the photoacclimation process in the fluorescence model in order to widen its applicability, such as predicting the evolution of microalgae culture over time periods of several days or even weeks. The preliminary results presented in Chapter 5 need extension. The aim should be to propose an empirical equation or a set of alternative empirical equations for each acclimation-dependent parameter. Ad hoc experiments could be specifically designed to an effective and reliable discrimination between alternatives. A calibration and validation against experimental data will be a crucial step of modelling development. Next, the dynamic modelling of photoacclimation could be introduced, a promising approach is the one proposed by Bernard (2011). The main originality of the model by Bernard (2011) is that it uses a conceptual variable, I^*, which is the irradiance at which the cells are photoacclimated. The variable I^* is in turn described by a differential equation which depends upon the growth rate, the average light intensity reaching the culture, and a time constant for the acclimation process. I^* is then linked to the pigment composition and to other acclimation dependent parameters, such as the total cross section of the photosystem. Eventually the vision is to integrate a fully validated model of photosynthesis within first-principles models describing the flow and light attenuation in large-scale microalgae culture systems as a mean to guide their design and operations. At this stage there will be

the need to design suitable experiments at varying light and to identify the most convenient modelling approaches and measurements to represent the system in a reasonably comprehensive way.

A second challenge could be to couple the hydrodynamic effect with light effect. A possible way to couple the light effect with the mixing effect will be to calculate the trajectories of the microalgae cells in the photobioreactor and to carry out a stochastic simulation considering the light profile at which the cells are exposed. The biomass productivity will be calculated from the average growth rate of the system. An alternative approach, more detailed but even more costly from a numerical point of view is to represent the system as a reacting system in which the microalgae growth is calculated along with the flow field and the variation of biomass concentration affects both the flow field and the light distribution. In the literature, a similar approach has been proposed by Hartmann et al. (2014) but the growth model utilised to describe the growth was quite simple. If a more complex growth model is used numerical issues might arise. The best compromise between numerical cost and accurate description of the system would have to be identified in order to assure the robustness of the approach.

A third interesting research direction is to study the effects of nutrients on growth and the possible interaction with light. This might be a very challenging task as there is still a lack of knowledge in the fundamental processes that regulates the effect of nutrients on the photosynthetic apparatus. Very recently experimental evidence that photosynthetic response is related to nitrogen concentration in rather sophisticated way has been published (Sforza et al., 2014). Those data could be exploited to formulate the first modelling hypothesis on the system. The various modelling hypotheses should be discriminated by means of MBDoE.

Appendix A

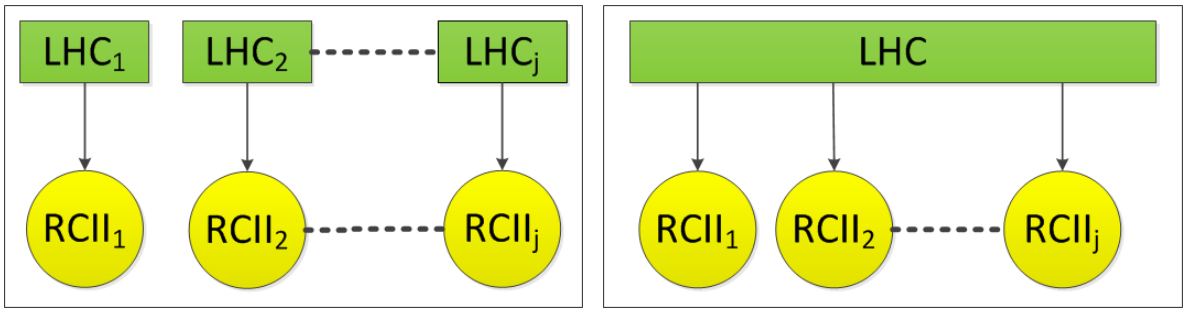
Fluorescence quantum yield

In this appendix we want to discuss the expression used for the calculation of Φ_f in Chapter 4. In particular we want to justify the utilisation of the harmonic mean to calculate the fluorescence quantum yield.

A.1 Alternative representation of the LHC-RC complex

As discussed in Chapter 2 two models are currently the most widely used in literature to represent the LHC-RC complex: the puddle and the lake model. A graphical representation of the two models is presented in Figure A.1.

In the puddle model every RC has its own antenna system. In Figure A.1(b) an array of j LHC-RC complexes is represented. Each complex will have a certain fluorescence yield $\Phi_{f,j}$, that depends on the oxidation state of the RC and on the NPQ activity. The fluorescence yield of the entire array is the algebraic mean of the j-th yields, as any of the j complexes is independent from the others. On the other hand, in a lake model configuration (Figure A.1(b)) the antenna system is shared by all the reaction centres and each reaction centre compete with the others for the excitation energy. The total fluorescence yield is therefore the harmonic mean



(a) puddle model (b) lake model

Figure A.1: *Graphical representation of the puddle (a) and the lake model (b) of the LHC-RC complex.*

of the j-th yields.

The utilisation of the harmonic mean can be justified as follows: the fluorescence flux from LHC is proportional to the average lifetime of the excited state of chlorophyll molecules, which in turn depends on the effectiveness of energy dissipation, according to the funnel scheme (see Section 1.4). If a reaction centre is closed its fluorescence yield is larger, because it cannot perform the photoproduction process. The higher yield can be seen as an higher resistance in the electron transfer from the LHC to the RC. Thus, since the resistances are placed in parallel, the equivalent resistance of all the reaction centres is the harmonic mean.

The correctness of utilising the harmonic mean is confirmed by the fact that the q_L parameter is consistent with the definition given by Kramer et al. (2004). In fact, if we consider the harmonic mean for the calculation of Φ_f, substituting the fluorescence fluxes expressions reported in Table 4.2 in the q_L definition, we obtain $q_L = A/(A+B)$ while with the algebraic mean the resulting expression for q_L is:

$$q_L = \frac{A}{A+B}\frac{(\Phi_f^A - \Phi_f^C)(A+B) + \Phi_f^C}{(\Phi_f^A - \Phi_f^C)A + (\Phi_f^B - \Phi_f^C)B) + \Phi_f^C} \tag{A.1}$$

which is different from the definition by Kramer et al. (2004).

Bibliography

Akaike, H., 1974. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* **19** (6), 716–723.

Alberton, A. L., Schwaab, M., Lobão, M. W. N., Pinto, J. C., 2012. Design of experiments for discrimination of rival models based on the expected number of eliminated models. *Chemical Engineering Science* **75**, 120–131.

Anning, T., MacIntyre, H. L., Pratt, S. M., Sammes, P. J., Gibb, S., Geider, R. J., 2000. Photoacclimation in the marine diatom skeletonema costatum. *Limnology and Oceanography* **45** (8), 1807–1817.

Aro, E.-M., Virgin, I., Andersson, B., 1993. Photoinhibition of photosystem ii. inactivation, protein damage and turnover. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1143** (2), 113–134.

Asprey, S., Macchietto, S., 2000. Statistical tools for optimal dynamic model building. *Computers & Chemical Engineering* **24** (2), 1261–1267.

Asprey, S., Macchietto, S., 2002. Designing robust optimal dynamic experiments. *Journal of Process Control* **12** (4), 545–556.

Baker, N. R., 2008. Chlorophyll fluorescence: A probe of photosynthesis in vivo. *Annual Review of Plant Biology* **59** (1), 89–113.

Baly, E., 1935. The kinetics of photosynthesis. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **117** (804), 218–239.

Barber, J., 2003. Photosystem ii: the engine of life. *Quarterly reviews of biophysics* **36** (01), 71–89.

Barber, J., Andersson, B., 1992. Too much of a good thing: light can be bad for photosynthesis. *Trends in biochemical sciences* **17** (2), 61–66.

Bard, Y., 1974. *Nonlinear parameter estimation*. Academic Press.

Béchet, Q., Shilton, A., Guieysse, B., 2013. Modeling the effects of light and temperature on algae growth: State of the art and critical assessment for productivity prediction during outdoor cultivation. *Biotechnology advances* **31** (8), 1648–1663.

Bellman, R., Åström, K. J., 1970. On structural identifiability. *Mathematical Biosciences* **7** (3), 329–339.

Bellu, G., Saccomani, M. P., Audoly, S., DAngiò, L., 2007. Daisy: A new software tool to test global identifiability of biological and physiological systems. *Computer methods and programs in biomedicine* **88** (1), 52–61.

Bernard, O., 2011. Hurdles and challenges for modelling and control of microalgae for CO_2 mitigation and biofuel production. *Journal of Process Control* **21** (10), 1378–1389.

Bernard, O., Gouzé, J.-L., 1999. Non-linear qualitative signal processing for biological systems: application to the algal growth in bioreactors. *Mathematical biosciences* **157** (1), 357–372.

Bernard, O., Rémond, B., 2012. Validation of a simple model accounting for light and temperature effect on microalgal growth. *Bioresource technology* **123**, 520–527.

Bernardi, A., Giarola, S., Bezzo, F., 2012. Optimizing the economics and the carbon and water footprints of bioethanol supply chains. *Biofuels, Bioproducts and Biorefining* **6** (6), 656–672.

Bernardi, A., Nikolaou, A., Meneghesso, A., Bezzo, F., Tomas, M., Chachuat, B., 2015a. A framework for the dynamic modelling of pi curves in microalgae. In: *PSE2015/ESCAPE25 conference*.

Bernardi, A., Nikolaou, A., Meneghesso, A., Bezzo, F., Tomas, M., Chachuat, B., 2015b. Using fluorescence measurements to model key phenomena in microalgae photosynthetic mechanisms. In: *Chemical Engineering Transaction*.

Bernardi, A., Perin, G., Sforza, E., Galvanin, F., Morosinotto, T., Bezzo, F., 2014. An identifiable state model to describe light intensity influence on microalgae growth. *Industrial & Engineering Chemistry Research* **53** (16), 6738–6749.

Biegler, L. T., Lang, Y.-d., Lin, W., 2014. Multi-scale optimization for process systems engineering. *Computers & Chemical Engineering* **60**, 17–30.

Bilger, W., Björkman, O., 1990. Role of the xanthophyll cycle in photoprotection elucidated by measurements of light-induced absorbance changes, fluorescence and photosynthesis in leaves of hedera canariensis. *Photosynthesis Research* **25** (3), 173–185.

Bonente, G., Pippa, S., Castellano, S., Bassi, R., Ballottari, M., 2012. Acclimation of chlamydomonas reinhardtii to different growth irradiances. *Journal of Biological Chemistry* **287** (8), 5833–5847.

Box, G. E., Hill, W., 1967. Discrimination among mechanistic models. *Technometrics* **9** (1), 57–71.

BP, 2014. *BP Statistical Review of World Energy June 2014*.

Burmaster, D. E., 1979. The unsteady continuous culture of phosphate-limited *Monochrysis lutheri* droop: Experimental and theoretical analysis. *Journal of experimental marine Biology and Ecology* **39** (2), 167–186.

Butler, W. L., 1978. Energy distribution in the photochemical apparatus of photosynthesis. *Annual Review of Plant Physiology* **29** (1), 345–378.

Chen, B. H., Asprey, S. P., 2003. On the design of optimally informative dynamic experiments for model discrimination in multiresponse nonlinear situations. *Industrial & Engineering Chemistry Research* **42** (7), 1379–1390.

Chinnasamy, S., Bhatnagar, A., Hunt, R. W., Das, K., 2010. Microalgae cultivation in a wastewater dominated by carpet mill effluents for biofuel applications. *Bioresource technology* **101** (9), 3097–3105.

Chisti, Y., 2007. Biodiesel from microalgae. *Biotechnology Advances* **25** (3), 294–306.

Demmig-Adams, B., Adams, W. W., 2002. Antioxidants in photosynthesis and human nutrition. *Science* **298** (5601), 2149–2153.

Dermoun, D., Chaumont, D., Thebault, J.-M., Dauta, A., 1992. Modelling of growth of *Porphyridium cruentum* in connection with two interdependent factors: Light and temperature. *Bioresource technology* **42** (2), 113–117.

Dochain, D., Vanrolleghem, P. A., Van Daele, M., 1995. Structural identifiability of biokinetic models of activated sludge respiration. *Water Research* **29** (11), 2571–2578.

Droop, M., 1968. Vitamin b12 and marine ecology. iv. the kinetics of uptake, growth and inhibition in monochrysis lutheri. *J. Mar. Biol. Assoc. UK* **48** (3), 689–733.

Droop, M., 1983. 25 years of algal growth kinetics a personal view. *Botanica marina* **26** (3), 99–112.

Duarte, P., 1995. A mechanistic model of the effects of light and temperature on algal primary productivity. *Ecological Modelling* **82** (2), 151–160.

Dubinsky, Z., Stambler, N., 2009. Photoacclimation processes in phytoplankton: mechanisms, consequences, and applications. *Aquat. Microb. Ecol* **56**, 163–176.

Dugdale, R., 1967. Nutrient limitation in the sea: dynamics, identification and significance. *Limnol. Oceanogr* **12** (4), 685–695.

Eberhard, S., Finazzi, G., Wollman, F.-A., 2008. The dynamics of photosynthesis. *Annual review of genetics* **42**, 463–515.

Eilers, P., Peeters, J., 1988. A model for the relationship between light intensity and the rate of photosynthesis in phytoplankton. *Ecological modelling* **42** (3), 199–215.

Falkowski, P. G., Dubinsky, Z., Wyman, K., 1985. Growth-irradiance relationships in phytoplankton. *Limnology and Oceanography* **30** (2), 311–321.

Falkowski, P. G., Greene, R., Kolber, Z., 1994. Light utilization and photoinhibition of photosynthesis in marine phytoplankton. In: *Photoinhibition of Photosynthesis*. Bio Scientific Publishers, Oxford, pp. 407–432.

Falkowski, P. G., Raven, J. A., 1997. *Aquatic photosynthesis*. Vol. 256. Blackwell Science Malden, MA.

Faugeras, B., Bernard, O., Sciandra, A., Lévy, M., et al., 2004. A mechanistic modelling and data assimilation approach to estimate the carbon/chlorophyll and carbon/nitrogen ratios in a coupled hydrodynamical-biological model. *Nonlinear Processes in Geophysics* **11**, 515–533.

Fisher, T., Minnaard, J., Dubinsky, Z., 1996. Photoacclimation in the marine alga nannochloropsis sp.(eustigmatophyte): a kinetic study. *Journal of plankton research* **18** (10), 1797–1818.

Flynn, K. J., 2001. A mechanistic model for describing dynamic multi-nutrient, light, temperature interactions in phytoplankton. *Journal of Plankton Research* **23** (9), 977–997.

Formighieri, C., Franck, F., Bassi, R., 2012. Regulation of the pigment optical density of an algal cell: filling the gap between photosynthetic productivity in the laboratory and in mass culture. *Journal of biotechnology* **162** (1), 115–123.

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science* **63** (19), 4846–4872.

Gaffron, H., Wohl, K., 1936. Zur theorie der assimilation. *Naturwissenschaften* **24** (6), 81–90.

Galvanin, F., Ballan, C. C., Barolo, M., Bezzo, F., 2013. A general model-based design of experiments approach to achieve practical identifiability of pharmacokinetic and pharmacodynamic models. *Journal of pharmacokinetics and pharmacodynamics* **40** (4), 451–467.

Galvanin, F., Barolo, M., Bezzo, F., 2009. Online model-based redesign of experiments for parameter estimation in dynamic systems. *Industrial & Engineering Chemistry Research* **48** (9), 4415–4427.

Galvanin, F., Barolo, M., Bezzo, F., Macchietto, S., 2010. A backoff strategy for model-based experiment design under parametric uncertainty. *AIChE journal* **56** (8), 2088–2102.

Galvanin, F., Macchietto, S., Bezzo, F., 2007. Model-based design of parallel experiments. *Industrial & engineering chemistry research* **46** (3), 871–882.

García-Camacho, F., Sánchez-Mirón, A., Molina-Grima, E., Camacho-Rubio, F., Merchuck, J., 2012. A mechanistic model of photosynthesis in microalgae including photoacclimation dynamics. *Journal of theoretical biology* **304**, 1–15.

Geider, R. J., MacIntyre, H. L., Kana, T. M., 1998. A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnology and Oceanography* **43** (4), 679–694.

Gentile, M.-P., Blanch, H. W., 2001. Physiology and xanthophyll cycle activity of nannochloropsis gaditana. *Biotechnology and bioengineering* **75** (1), 1–12.

Genty, B., Briantais, J.-M., Baker, N. R., 1989. The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta (BBA) - General Subjects* **990** (1), 87 – 92.

Gerbens-Leenes, W., Hoekstra, A. Y., van der Meer, T. H., 2009. The water footprint of bioenergy. *Proceedings of the National Academy of Sciences* **106** (25), 10219–10223.

Gerotto, C., Alboresi, A., Giacometti, G. M., Bassi, R., Morosinotto, T., 2011. Role of psbs and lhcsr in physcomitrella patens acclimation to high light and low temperature. *Plant, cell & environment* **34** (6), 922–932.

Haag, A. L., 2007. Algae bloom again. *Nature* **447**, 520–521.

Han, B. P., 2001. Photosynthesis-irradiance response at physiological level: A mechanistic model. *Journal of Theoretical Biology* **213**, 121–127.

Han, B. P., 2002. A mechanistic model of algal photoinhibition induced by photodamage to photosystem-II. *Journal of Theoretical Biology* **214** (4), 519–27.

Han, B.-P., Virtanen, M., Koponen, J., Straškraba, M., 2000a. Effect of photoinhibition on algal photosynthesis: a dynamic model. *Journal of Plankton Research* **22** (5), 865–885.

Han, B.-P., Virtanen, M., Koponen, J., Straškraba, M., 2000b. Effect of photoinhibition on algal photosynthesis: a dynamic model. *Journal of Plankton Research* **22** (5), 865–885.

Hannon, M., Gimpel, J., Tran, M., Rasala, B., Mayfield, S., 2010. Biofuels from algae: challenges and potential. *Biofuels* **1** (5), 763–784.

Hartmann, P., Béchet, Q., Bernard, O., 2014. The effect of photosynthesis time scales on microalgae productivity. *Bioprocess and biosystems engineering* **37** (1), 17–25.

Hartmann, P., Nikolaou, A., Chachuat, B., Bernard, O., 2013. A dynamic model coupling photoacclimation and photoinhibition in microalgae. In: *European Control Conference (ECC'13)*. July 17-19, 2013, Zürich, Switzerland.

Hawkins, D. M., 2004. The problem of overfitting. *Journal of chemical information and computer sciences* **44** (1), 1–12.

Horton, P., Hague, A., 1988. Studies on the induction of chlorophyll fluorescence in isolated barley protoplasts. iv. resolution of non-photochemical quenching. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **932**, 107–115.

Horton, P., Ruban, A., 2005. Molecular design of the photosystem ii light-harvesting antenna: photosynthesis and photoprotection. *Journal of experimental botany* **56** (411), 365–373.

Hubbert, M. K., 1949. *Energy from fossil fuels*.

Huot, Y., Babin, M., 2010. Overview of fluorescence protocols: theory, basic concepts, and practice. In: *Chlorophyll a Fluorescence in Aquatic Sciences: Methods and Applications*. Springer, pp. 31–74.

IEA, 2011. *World energy outlook 2011*. International Energy Agency: Paris, France.

IEA, 2014. *World energy outlook 2014*. International Energy Agency: Paris, France.

IPCC, 2014. *Climate Change 2014 synthesis report*. Intergovernal Panel on Climate Change: Geneva, Swizerland.

Jansen, M. A., Mattoo, A. K., Edelman, M., 1999. D1-d2 protein degradation in the chloroplast. *European journal of biochemistry* **260** (2), 527–532.

Jassby, A. D., Platt, T., 1976. *Mathematical formulation of the relationship between photosynthesis and light for phytoplankton*.

Joliot, P., Joliot, A., 2006. Cyclic electron flow in c3 plants. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1757** (5), 362–368.

Jorquera, O., Kiperstok, A., Sales, E. A., Embiruçu, M., Ghirardi, M. L., 2010. Comparative energy life-cycle analyses of microalgal biomass production in open ponds and photobioreactors. *Bioresource technology* **101** (4), 1406–1413.

Kaufmann, R. K., 2011. The role of market fundamentals and speculation in recent price changes for crude oil. *Energy Policy* **39** (1), 105–115.

Kiefer, J., 1959. Optimum experimental designs. *Journal of the Royal Statistical Society. Series B (Methodological)* **21**, 272–319.

Kirst, H., Melis, A., 2014a. The chloroplast signal recognition particle (cpsrp) pathway as a tool to minimize chlorophyll antenna size and maximize photosynthetic productivity. *Biotechnology advances* **32** (1), 66–72.

Kitajima, M., Butler, W., 1975. Quenching of chlorophyll fluorescence and primary photochemistry in chloroplasts by dibromothymoquinone. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **376** (1), 105–115.

Kok, B., 1956. On the inhibition of photosynthesis by intense light. *Biochimica et biophysica acta* **21** (2), 234–244.

Kolber, Z., Falkowski, P. G., 1993. Use of active fluorescence to estimate phytoplankton photosynthesis in situ. *Limnology and Oceanography* **38** (8), 1646–1665.

Kong, Q.-x., Li, L., Martinez, B., Chen, P., Ruan, R., 2010. Culture of microalgae chlamydomonas reinhardtii in wastewater for biomass feedstock production. *Applied biochemistry and Biotechnology* **160** (1), 9–18.

Kramer, D., Johnson, G., Kiirats, O., Edwards, G., 2004. New fluorescence parameters for the determination of q(a) redox state and excitation energy fluxes. *Photosynthesis Research* **79** (2), 1209–218.

Krichnavaruk, S., Shotipruk, A., Goto, M., Pavasant, P., 2008. Supercritical carbon dioxide extraction of astaxanthin from *Haematococcus pluvialis* with vegetable oils as co-solvent. *Bioresource Technology* **99** (13), 5556–5560.

Kroon, B., Thoms, S., 2006. From electron to biomass: A mechanistic model to describe phytoplankton photosynthesis and steady-state growth rates1. *Journal of phycology* **42** (3), 593–609.

Lavergne, J., Trissl, H.-W., 1995. Theory of fluorescence induction in photosystem ii: derivation of analytical expressions in a model including exciton-radical-pair equilibrium and restricted energy transfer between photosynthetic units. *Biophysical journal* **68** (6), 2474–2492.

Lazr, D., 1999. Chlorophyll a fluorescence induction. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1412** (1), 1 – 28.

Li, X., Xu, H., Wu, Q., 2007. Large-scale biodiesel production from microalga chlorella protothecoides through heterotrophic cultivation in bioreactors. *Biotechnology and bioengineering* **98** (4), 764–771.

Li, Y., Horsman, M., Wu, N., Lan, C. Q., Dubois-Calero, N., 2008. Biofuels from microalgae. *Biotechnology progress* **24** (4), 815–820.

Li, Z., Wakao, S., Fischer, B. B., Niyogi, K. K., 2009. Sensing and responding to excess light. *Annual review of plant biology* **60**, 239–260.

Ljung, L., Glad, T., 1994. On global identifiability for arbitrary model parametrizations. *Automatica* **30** (2), 265–276.

Loebl, M., Cockshutt, A. M., Campbell, D. A., Finkel, Z. V., 2010. Physiological basis for high resistance to photoinhibition under nitrogen depletion in emiliania huxleyi. *Limnology and Oceanography* **55** (5), 2150.

Long, S., Humphries, S., Falkowski, P. G., 1994. Photoinhibition of photosynthesis in nature. *Annual review of plant biology* **45** (1), 633–662.

MacIntyre, H. L., Kana, T. M., Anning, T., Geider, R. J., 2002. Photoacclimation of photosynthesis irradiance response curves and photosynthetic pigments in microalgae and cyanobacteria1. *Journal of Phycology* **38** (1), 17–38.

Marshall, H. L., Geider, R. J., Flynn, K. J., 2000. A mechanistic model of photoinhibition. *New phytologist* **145** (2), 347–359.

Mata, T. M., Martins, A. A., Caetano, N. S., 2010. Microalgae for biodiesel production and other applications: a review. *Renewable and Sustainable Energy Reviews* **14** (1), 217–232.

Maxwell, K., Johnson, G. N., 2000. Chlorophyll fluorescence - a practical guide. *Journal of experimental botany* **51** (345), 659–668.

Meneghesso, A., Bernardi, A., Nikolaou, A., Perin, G., Chachuat, B., Bezzo, F., Morosinotto, T., submitted. Computational models to improve light use efficiency in microalgae culture. *Journal of Biotechnology*.

Meshkat, N., Eisenberg, M., DiStefano III, J. J., 2009. An algorithm for finding globally identifiable parameter combinations of nonlinear ode models using gröbner bases. *Mathematical biosciences* **222** (2), 61–72.

Miao, H., Xia, X., Perelson, A. S., Wu, H., 2011. On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM review* **53** (1), 3–39.

Miyao, M., 1994. Involvement of active oxygen species in degradation of the d1 protein under strong illumination in isolated subcomplexes of photosystem ii. *Biochemistry* **33** (32), 9722–9730.

Moody, J. W., McGinty, C. M., Quinn, J. C., 2014. Global evaluation of biofuel potential from microalgae. *Proceedings of the National Academy of Sciences*, 201321652.

Müller, P., Li, X.-P., Niyogi, K. K., 2001. Non-photochemical quenching. a response to excess light energy. *Plant Physiology* **125** (4), 1558–1566.

Mutanda, T., Ramesh, D., Karthikeyan, S., Kumari, S., Anandraj, A., Bux, F., 2011. Bioprospecting for hyper-lipid producing microalgal strains for sustainable biofuel production. *Bioresource Technology* **102** (1), 57–70.

Nauha, E. K., Alopaeus, V., 2013. Modeling method for combining fluid dynamics and algal growth in a bubble column photobioreactor. *Chemical Engineering Journal* **229**, 559–568.

Nigam, P. S., Singh, A., 2011. Production of liquid biofuels from renewable resources. *Progress in Energy and Combustion Science* **37** (1), 52–68.

Nikolaou, A., Bernardi, A., Bezzo, F., Tomas, M., Chachuat, B., 2014a. A dynamic model of photoproduction, photoregulation and photoihibition in microalgae using chlorophyll fluorescence. In: *Preprints of the IFAC 19th World Congress*. Cape Town, South Africa. August 24-29, pp. 4370–4375.

Nikolaou, A., Bernardi, A., Meneghesso, A., Bezzo, F., Morosinotto, T., Chachuat, B., 2015. A model of chlorophyll fluorescence in microalgae integrating photoproduction, photoinhibition and photoregulation. *Journal of Biotechnology* **194**, 91 – 99.

Nixon, P. J., Michoux, F., Yu, J., Boehm, M., Komenda, J., 2010. Recent advances in understanding the assembly and repair of photosystem ii. *Annals of botany* **106**, 1–16.

Norberg, J., 2004. Biodiversity and ecosystem functioning: a complex adaptive systems approach. *Limnology and Oceanography* **49** (4), 1269–1277.

Oxborough, K., Baker, N. R., 1997. Resolving chlorophyll a fluorescence images of photosynthetic efficiency into photochemical and non-photochemical components–calculation of qp and fv-/fm-; without measuring fo. *Photosynthesis research* **54** (2), 135–142.

Oxborough, K., Baker, N. R., 2000. An evaluation of the potential triggers of photoinactivation of photosystem ii in the context of a stern–volmer model for downregulation and the reversible radical pair equilibrium model. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **355** (1402), 1489–1498.

Pahlow, M., 2005. Linking chlorophyll-nutrient dynamics to the redfield n: C ratio with a model of optimal phytoplankton growth. *Marine Ecology Progress Series* **287**, 33–43.

Papadakis, I. A., Kotzabasis, K., Lika, K., 2012. Modeling the dynamic modulation of light energy in photosynthetic algae. *Journal of theoretical biology* **300**, 254–264.

Park, S., Khamai, P., Garcia-Cerdan, J. G., Melis, A., 2007. Rep27, a tetratricopeptide repeat nuclear-encoded and chloroplast-localized protein, functions in d1/32-kd reaction center protein turnover and photosystem ii repair from photodamage. *Plant physiology* **143** (4), 1547–1560.

Peers, G., Truong, T. B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A. R., Hippler, M., Niyogi, K. K., 2009. An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462** (7272), 518–521.

Pinnola, A., DallOsto, L., Gerotto, C., Morosinotto, T., Bassi, R., Alboresi, A., 2013. Zeaxanthin binds to light-harvesting complex stress-related protein to enhance nonphotochemical quenching in physcomitrella patens. *The Plant Cell Online* **25** (9), 3519–3534.

Pittman, J. K., Dean, A. P., Osundeko, O., 2011. The potential of sustainable algal biofuel production using wastewater resources. *Bioresource Technology* **102** (1), 17–25.

Pohjanpalo, H., 1978. System identifiability based on the power series expansion of the solution. *Mathematical biosciences* **41** (1), 21–33.

Process System Enterprise, a, a., 2012. *gPROMS model validation guide (v. 3.6)*. London: Process System Enterprise Ltd.

Quinn, J. C., Catton, K., Wagner, N., Bradley, T. H., 2012. Current large-scale us biofuel potential from microalgae cultivated in photobioreactors. *BioEnergy Research* **5** (1), 49–60.

Radakovits, R., Jinkerson, R. E., Darzins, A., Posewitz, M. C., 2010. Genetic engineering of algae for enhanced biofuel production. *Eukaryotic cell* **9** (4), 486–501.

Rees, D., Noctor, G., Horton, P., 1990. The effect of high-energy-state excitation quenching on maximum and dark level chlorophyll fluorescence yield. *Photosynthesis research* **25** (3), 199–211.

Richmond, A., 2008. *Handbook of microalgal culture: biotechnology and applied phycology*. John Wiley & Sons.

Riley, G. A., 1946. Factors controlling phytoplankton populations on georges bank. *Journal of marine Research* **6** (1), 54–73.

Roháček, K., Barták, M., 1999a. Technique of the modulated chlorophyll fluorescence: basic concepts, useful parameters, and some applications. *Photosynthetica* **37** (3), 339–363.

Roháček, K., Barták, M., 1999b. Technique of the modulated chlorophyll fluorescence: basic concepts, useful parameters, and some applications. *Photosynthetica* **37** (3), 339–363.

Ross, O. N., Geider, R. J., 2009. New cell-based model of photosynthesis and photo-acclimation: accumulation and mobilisation of energy reserves in phytoplankton. *Marine Ecology Progress Series* **383**, 53–71.

Ross, O. N., Moore, C. M., Suggett, D. J., MacIntyre, H. L., Geider, R. J., 2008. A model of photosynthesis and photo-protection based on reaction center damage and repair. *Limnology and Oceanography* **53** (5), 1835.

Rosso, L., Lobry, J., Flandrois, J., 1993. An unexpected correlation between cardinal temperatures of microbial growth highlighted by a new model. *Journal of Theoretical Biology* **162** (4), 447–463.

Ruban, A. V., Murchie, E. H., 2012. Assessing the photoprotective effectiveness of non-photochemical chlorophyll fluorescence quenching: a new approach. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1817** (7), 977–982.

Rubio, F. C., Camacho, F. G., Sevilla, J., Chisti, Y., Grima, E. M., 2003. A mechanistic model of photosynthesis in microalgae. *Biotechnology and bioengineering* **81** (4), 459–473.

Rubio, F. C., Fernandez, F., Perez, J., Camacho, F. G., Grima, E. M., 1999. Prediction of dissolved oxygen and carbon dioxide concentration profiles in tubular photobioreactors for microalgal culture. *Biotechnology and Bioengineering* **62** (1), 71–86.

Ryther, J. H., 1955. The ratio of photosynthesis to respiration in marine plankton algae and its effect upon the measurement of productivity. *Deep Sea Research (1953)* **2** (2), 134–139.

Saccomani, M. P., Audoly, S., D'Angiò, L., 2003. Parameter identifiability of nonlinear systems: the role of initial conditions. *Automatica* **39** (4), 619–632.

Schenk, P. M., Thomas-Hall, S. R., Stephens, E., Marx, U. C., Mussgnug, J. H., Posten, C., Kruse, O., Hankamer, B., 2008. Second generation biofuels: high-efficiency microalgae for biodiesel production. *Bioenergy Research* **1** (1), 20–43.

Schreiber, U., 1986. Detection of rapid induction kinetics with a new type of high-frequency modulated chlorophyll fluorometer. In: *Current topics in photosynthesis*. Springer, pp. 259–270.

Sforza, E., Simionato, D., Giacometti, G. M., Bertucco, A., Morosinotto, T., 2012. Adjusted light and dark cycles can optimize photosynthetic efficiency in algae growing in photobioreactors. *PloS one* **7** (6), e38975.

Sforza, E., Urbani, S., Bertucco, A., 2014. Evaluation of maintenance energy requirements in the cultivation of scenedesmus obliquus: effect of light intensity and regime. *Journal of Applied Phycology*, 1–10.

Sheehan, J., Dunahay, T., Benemann, J., Roessler, P., 1998. A Look Back at the U.S. Department of Energy's Aquatic Species Program – Biodiesel from Algae. Tech. rep., U.S. Department of Energy.

Sialve, B., Bernet, N., Bernard, O., 2009. Anaerobic digestion of microalgae as a necessary step to make microalgal biodiesel sustainable. *Biotechnology advances* **27** (4), 409–416.

Simionato, D., Basso, S., Giacometti, G. M., Morosinotto, T., 2013a. Optimization of light use efficiency for biofuel production in algae. *Biophysical chemistry* **182**, 71–78.

Simionato, D., Block, M. A., La Rocca, N., Jouhet, J., Maréchal, E., Finazzi, G., Morosinotto, T., 2013b. The response of nannochloropsis gaditana to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. *Eukaryotic cell* **12** (5), 665–676.

Simionato, D., Sforza, E., Corteggiani Carpinelli, E., Bertucco, A., Giacometti, G. M., Morosinotto, T., 2011. Acclimation of *Nannochloropsis gaditana* to different illumination regimes: Effects on lipids accumulation. *Bioresource technology* **102** (10), 6026–6032.

Sims, R. E., Mabee, W., Saddler, J. N., Taylor, M., 2010. An overview of second generation biofuel technologies. *Bioresource technology* **101** (6), 1570–1580.

Soeder, C. J., 1980. Massive cultivation of microalgae: results and prospects. *Hydrobiologia* **72** (1-2), 197–209.

Spolaore, P., Joannis-Cassan, C., Duran, E., Isambert, A., 2006. Commercial applications of microalgae. *Journal of bioscience and bioengineering* **101** (2), 87–96.

Stern, N. H., Britain, G., Treasury, H., 2006. *Stern Review: The economics of climate change*. Vol. 30. HM treasury London.

Stewart, W., Shon, Y., Box, G., 1998. Discrimination and goodness of fit of multiresponse mechanistic models. *AIChE journal* **44** (6), 1404–1412.

Stewart, W. E., Henson, T. L., Box, G. E., 1996. Model discrimination and criticism with single-response data. *AIChE journal* **42** (11), 3055–3062.

Suggett, D. J., Oxborough, K., Baker, N. R., MacIntyre, H. L., Kana, T. M., Geider, R. J., 2003. Fast repetition rate and pulse amplitude modulation chlorophyll a fluorescence measurements for assessment of photosynthetic electron transport in marine phytoplankton. *European Journal of Phycology* **38** (4), 371–384.

Taiz, L., Zeiger, E., 2010. Plant physiology. *Sunderland, MA: Sinauer Associates*.

Vance, C. P., 2001. Symbiotic nitrogen fixation and phosphorus acquisition. plant nutrition in a world of declining renewable resources. *Plant physiology* **127** (2), 390–397.

Vass, I., 2012. Molecular mechanisms of photodamage in the photosystem ii complex. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1817** (1), 209–217.

Vatcheva, I., De Jong, H., Bernard, O., Mars, N. J., 2006. Experiment selection for the discrimination of semi-quantitative models of dynamical systems. *Artificial Intelligence* **170** (4), 472–506.

Vollenweider, R. A., 1966. Calculation models of photosynthesis-depth curves and some implications regarding day rate estimates in primary production measurements. *Primary productivity in aquatic environments* **18**, 425.

Walter, E., Lecourtier, Y., 1981. Unidentifiable compartmental models: What to do? *Mathematical biosciences* **56** (1), 1–25.

Walter, E., Pronzato, L., 1997. Identification of parametric models. *Communications and Control Engineering*.

Walters, R. G., 2005. Towards an understanding of photosynthetic acclimation. *Journal of experimental botany* **56** (411), 435–447.

Williams, P. J. l. B., Laurens, L. M. L., 2010. Microalgae as biodiesel & biomass feedstocks: Review & analysis of the biochemistry, energetics & economics. *Energy & Environmental Science* **3** (5), 554–590.

Wobbe, L., Remacle, C., 2014. Improving the sunlight-to-biomass conversion efficiency in microalgal biofactories. *Journal of biotechnology*.

Wu, X., Merchuk, J. C., 2001. A model integrating fluid dynamics in photosynthesis and photoinhibition processes. *Chemical Engineering Science* **56** (11), 3527–3538.

Zonneveld, C., 1998. A cell-based model for the chlorophyll *a* to carbon ratio in phytoplankton. *Ecological Modelling* **113** (1), 55–70.

Zullo, L. C., 1991. *Computer aided design of experiments. An engineering approach*. The University of London, London (U.K.).