UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA INDUSTRIALE

CORSO DI DOTTORATO DI RICERCA IN: INGEGNERIA INDUSTRIALE
CURRICULUM: INGEGNERIA CHIMICA, DEI MATERIALI E MECCANICA
CICLO: XXIX

AN INVESTIGATION ON MICROALGAE GROWTH AT DIFFERENT SCALES: FROM PHOTOSYNTHETIC MECHANISMS MODELLING TO OPERATION OPTIMISATION IN OPEN POND CULTIVATION SYSTEMS.

DIRETTORE DELLA SCUOLA: CH.MO PROF. PAOLO COLOMBO
COORDINATORE D'INDIRIZZO: CH.MO PROF. GIOVANNI MENEGHETTI
SUPERVISORE : CH.MO PROF. FABRIZIO BEZZO

DOTTORANDO: RICCARDO DE LUCA

Foreword

The research project presented in this Thesis has involved the financial and intellectual support of many people, to whom the author is most grateful. The research activity reported in the Thesis has been mainly 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 also carried out at INRIA-Biocore, Sophia Antipolis (FR), under the supervision of Prof. Olivier Bernard. The realisation of this study has been made possible thanks to the financial support of 'VII UE Program IDEAS-ERC 2012' for the project 'Biotechnological optimization of light use efficiency in algae photobioreactors'. The full list of publications and presentations made during the PhD period is reported below.

Articles published in international journals

Galvanin F., R. De-Luca, G. Ferrentino, M. Barolo, S. Spilimbergo, F. Bezzo (2014). Bacterial inactivation on solid food matrices through supercritical CO2: a correlative study. *J. Food Eng.*, 120:146-157;
De-Luca R., F. Galvanin, F. Bezzo. (2016). A methodology for direct exploitation of available information in the online model-based redesign of experiments. *Comput. Chem Eng.*, 91:195-205;
De-Luca R., Q. Béchet, F. Bezzo, O. Bernard (2016). Optimal management of raceways using weather forecasts. *IFAC-PapersOnLine*, 49(7):1062-1067.

Articles accepted in international journals

De-Luca R., F. Bezzo, Q. Béchet, O. Bernard (2017). Exploiting meteorological forecasts for the optimal operation of algal ponds. In: *Journal of Process Control.* doi: 10.1016/j.jprocont.2017.03.010

Articles submitted in international journals

De-Luca R., F. Bezzo, Q. Béchet, O. Bernard (2017). Meteorological data-based optimal control strategy for microalgae cultivation in open pond systems. In: *Bioresource Technology.*

Conference proceedings

Ferrari, M., L. Bosa, R. De-Luca, M. Barolo, C. F. Zambon, V. Pengo, R. Padrini, F. Bezzo (2016). A pharmacokinetic-pharmacodynamic model for individualisation of an oral anticoagulation therapy. In: *Computer-Aided Chemical Engineering 38, 26th European Symposium on Computer Aided Process Engineering (Z. Kravanja and M. Bogataj, Ed.), Elsevier, Amsterdam (The Netherlands)*, 2313-2318.

Oral presentations

De-Luca, R., F. Galvanin, F. Bezzo (2015). A framework for direct exploitation of available information in the online model-based redesign of experiments. In: *12th International Symposium on Process Systems Engineering and 25th European Symposium in Computer Aided Process Engineering. (31 May - 4 June 2015, Copenhagen, Denmark)*;
De-Luca, R., Q. Béchet, O. Bernard, F. Bezzo (2016). Optimal management of raceways using meteorological forecasts. In: *11th IFAC Symposium on Dynamics and Control of Process Systems. (06 June - 08 June 2016, Trondheim, Norway)*;
De-Luca, R., Q. Béchet, O. Bernard, F. Bezzo (2016). Utilizzo di dati meteorologici per l'ottimizzazione di microalghe in sistemi open pond. In: *GRICU 2016 (Chemical Engineering for sustainable production of energy and fne chemicals). (12 September - 17 September 2016, Anacapri, Italy).*

Abstract

Microalgae processing represents one of the most promising new technologies for sustainable production of a wide range of commodities and value-added products, including cosmetics, pharmaceuticals and nutraceuticals. Morevoer, at a larger time horizon, microalgae are expected to contribute for fossil carbon replacement with renewable carbon, especially for supplying green chemicals and liquid biofuel in the transport sector. Nevertheless, much research is still needed in order to make this potential new energy source a practically and economically feasible technology, since all the existing technological assessments are based on specific assumptions or gross estimates of productivity, derived by extrapolation of laboratory-scale data.

The development of reliable mathematical models predicting both the behavior of large-scale outdoor microalgae culture and the underlying multiple time-scale biophysical and chemical processes is therefore necessary. These models are valuable tools to support both system design and operation optimization, with consequent potential increase of the process profitability. This Thesis aims at investigating the complex behavior of microalgae growth by following two main approaches.

The first objective was to extend an existing growth model of marine water alga *Nannochloropsis Salina* describing photosynthetic efficiency through chlorophyll fluorescence dynamics. This micro-scale model integrates photoproduction, photoregulation and photoinhibition processes in a semi-mechanistic way, but it is limited to the description of the most significant photosystem (PSII). The proposed model extension aims at describing the complete electron transport, together with the dynamics of each protein complex involved in the photosynthetic process, through absorbance data-based calibration/validation. The results show that the calibrated model is capable of accurate quantitative predictions of the photosynthetic transport chain paths under a wide range of transient light conditions.

The second contribution objective was to develop a macro-scale model for *Chlorella Vulgaris* cultivation in open pond systems by coupling existing growth/temperature sub-models with real meteorological data. The utilization of this dynamic model will underline the benefits of model building activities on practical process optimization, since a reduced set of 'rules of thumb' was extracted by different simulations done at different weather conditions. The proposed optimization strategy significantly increased productivity compared to standard operation at constant dilution rate and pond depth, by up to a factor 2.2. Furthermore, a deeper insight into optimal operation in case of inaccurate forecasts has been developed and discussed. The different strategies proposed can guarantee both high productivity and feasible operation in case of inaccurate weather forecasts. The resulting control strategies, despite the high amount of water required, can prevent culture death conditions due to unpredicted high temperatures.

Riassunto

L'incremento dell'instabilità dei prezzi petroliferi, la volontà di incentivare l'utilizzo di risorse rinnovabili e la necessità di affrontare problematiche relative al riscaldamento globale sono alcuni dei motivi che hanno incrementato lo sforzo della comunità scientifica al fine di utilizzare le microalghe a scopo energetico. Queste ultime, infatti, presentano un'alta velocità di crescita e sono in grado di contenere ricche quantià di olio; oltre a questo possono essere coltivate in zone dedicate, riducendo l'utilizzo dei terreni per unità di biomassa prodotta, possono essere utilizzate per mitigare l'impatto ambientale dovuto a emissioni di CO_2 (essendo questa fissata dalle cellule durante il processo fotosintetico) e, necessitando dei nutrienti presenti in acque reflue inquinate, possono essere utilizzate anche a fini depurativi.

Nonostante vi sia un alto potenziale teorico per l'utilizzo di microalghe come materia prima per la produzione di biocarburanti e prodotti ad alto valore aggiunto (generalmente destinati al settore cosmetico, farmacologico e nutraceutico), la produzione su larga scala risulta attualmente non competitiva in termini economici. Le previsioni più ottimistiche, infatti, sono basate su estrapolazioni di dati ottenuti in laboratorio in condizioni controllate, che risultano notevolmente differenti da quelle che si verificano su larga scala. Oltre a questo, le valutazioni tecno-economiche basate su estrapolazioni di dati di laboratorio dipendono fortemente dalla metodologia applicata. La disponibilità di modelli meccanicistici in grado di prevedere il comportamento delle colture microalgali in sistemi di coltivazione industriali risulta quindi di primaria importanza per progettare, simulare e ottimizzare i processi di produzione.

Il primo obiettivo di questa Tesi è stato quello di estendere un modello di microscala in grado di desrivere la crescita della microalga marina *Nannochloropsis Salina* tramite la rappresentazione dei meccanismi di fluorescenza causati dall'attività fotosintetica. Questo modello semi-meccanicistico descrive i meccanismi di fotoproduzione, fotoregolazione e fotoinibizione attraverso la rappresentazione della dinamica del fotosistema (PSII), complesso proteico coinvolto nel processo fotosintetico. Il lavoro presentato in questa Tesi consiste nell'estensione del suddetto modello tramite la descrizione dell'intero processo di trasporto elettronico a livello della membrana tilacoidale e delle dinamiche di ossidoriduzione dei complessi proteici coinvolti nel processo fotosintetico. Il modello è stato calibrato e validato grazie a dati di assorbanza ricavati tramite spettrometro di tipo Joliot. I risultati ottenuti dimostrano come il modello sia in grado fornire previsioni quantitative accurate del trasporto fotosintetico per una vasta gamma di condizioni di luce.

Il secondo obiettivo è stato quello di sviluppare un modello di macroscala basato su previsioni meteorologiche per la coltivazione di *Chlorella Vulgaris* in sistemi a vasca aperta. L'utilizzo di questo modello dinamico sottolinerà i benefici derivanti dall'attività di modellazione ai fini dell'ottimizzazione di processo. Infatti la strategia di ottimizzazione proposta permetterà di ottenere un significativo aumento di produttività (fino a un fattore 2.2 per le stagioni più calde e soleggiate) rispetto a quanto può essere ottenuto applicando le tradizionali strategie di coltivazione (a tempo di permanenza e volume costante). L'attività di ricerca ha inoltre individuato e paragonato differenti strategie di controllo che permettono di garantire la fattibilità del processo nel caso in cui i dati meteorologici non risultino accurati. Gli approcci studiati permettono di evitare condizioni critiche a seguito di ottimizzazioni basate su previsioni meteorologici inaccurati, richiedendo un maggiore consumo d'acqua per la regolazione del sistema, ma riuscendo in tal modo a garantire elevati valori di produttività.

Il lavoro di Tesi è organizzato secondo il seguente schema concettuale.
Nel Capitolo 1 viene introdotta una panoramica riguardante l'attuale utilizzo di microlalghe come risorsa rinnovabile e una breve descrizione relativa agli attuali processi di coltivazione; succcessivamente vengono illustrati i principali approcci modellistici usati per descrivere i meccanismi di crescita delle microalghe. Nella parte finale del capitolo si andranno a evidenziare gli obiettivi di ricerca di questo lavoro di Tesi.
Il Capitolo 2 presenta un'estensione del modello semi-meccanicistico proposto da Bernardi et al. (2016). Questo modello a stati descrive i principali processi biologici che agiscono sul fotosistema PSII presente nei cloroplasti delle microalghe: tali processi ricoprono scale temporali che vanno dai millisecondi (fenomeni di fotoproduzione) a minuti/ore/giorni (*Non-Photochemical Quenching*, fotoinibizione, fotoacclimatazione). Il modello è stato sviluppato e convalidato utilizzando esperimenti di flourescenza. L'estensione del modello proposta in questo lavoro di Tesi mira a descrivere le dinamiche dei fondamentali complessi proteici coinvolti nella catena di trasporto elettronico facendo uso di dati assorbanza della luce ottenuti tramite spettrometro di tipo Joliot (JTS-10).
Il Capitolo 3 presenta la descrizione dettagliata di un nuovo modello dinamico proposto per descrivere la produttività microalgale in vasche aperte; questo modello è il risultato della combinazione di modelli esistenti in grado di prevedere sia l'impatto della temperatura sul sistema di coltivazione (e sui parametri di crescita del ceppo algale coltivato) sia l'impatto della distribuzione della luce sulla crescita delle microalghe. Inoltre, la variabilità delle condizioni climatiche a cui avviene la fase di coltivazione è rappresentata attraverso l'implementazione di previsioni meteorologiche reali. Il modello proposto è stato utilizzato per svolgere differenti attività di ottimizzazione al fine di verificare i vantaggi derivati dalla manipolazione continua dei flussi in ingresso e in uscita dal reattore rispetto alle tradizionali modalità di controllo della vasca.
Nel Capitolo 4 viene fornita un'analisa dettagliata delle strategie di ottimizzazione ottenute per diverse stagioni dell'anno, derivando un numero ridotto di 'rules of thumb' utilizzabili in future applicazioni pratiche. L'analisi è stata condotta sia per il caso ideale in cui si

assume di avere a disposizione dati meteorologici perfetti, sia per il caso reale in cui i dati meteorologici risultano affetti da incertezza.

Il Capitolo 5 presenta uno studio di fattibilità condotto al fine di prevenire la perdita di produttività che può derivare da ottimizzazioni basate su previsioni meteorologiche non accurate. Infatti, l'eventuale strategia di controllo ottenuta tramite ottimizzazione con dati inaccurati può portare al raggiungimento di condizioni critiche per la crescita (e.g. temperature del sistema di coltivazione eccessivamente alte con conseguente rischio di morte cellulare). Sono stati quindi proposti due differenti approcci di ottimizzazione al fine di garantire la fattibilità della coltivazione anche in assenza di dati precisi. I due metodi verranno infine confrontati in termini di produttività, fabbisogno idrico e relativi costi di processo. Il Capitolo 6 riassume i principali risultati raggiunti in questo lavoro di Tesi e propone alcuni suggerimenti per futuri lavori di ricerca.

In Appendice A viene discusso un metodo alternativo di progettazione sperimentale basata su modello (MBDoE). L'approccio proposto è fondato sul concetto di riprogettazione sperimentale online al fine di sfruttare l'aumento progressivo dell'informazione derivante dall'esecuzione dell'esperimento stesso. L'efficacia di questa tecnica, proposta in particolare per casi ad alto mismatch parametrico, dimostrata tramite due casi studio non strettamente inerenti all'argomento della Tesi. Tuttavia, si riportano i risultati ottenuti a causa del loro potenziale per attività future di calibrazione/validazione di modelli di crescita microalgale.

Contents

Acronyms

AID	adaptive input design.
ATP	adenosine triphosphate.
CAPEX	capital cost.
CEF	cyclic electron flow.
CSP	concentrated solar photovoltaic.
Cytb6f	cytochrome b6f complex.
Cytc6	cytochrome c6 complex.
DAEs	differential and algebraic equations.
DBMIB	2,5-Dibrom-3-methyl-6-isopropyl-p-benzochinon.
DCMU	3-(3,4-dichlorophenyl)-1,1-dimethylurea.
ECMWF	European Centre for Medium-Range Weather Forecast.
EGE	enhanced geothermal energy.
Fd	ferrodoxin.
GHG	greenhouse gas.
HTR	hydraulic retention time.
IDRO	information-driven redesign optimization.
LCA	life cycle analysis.
LEF	linear electron flow.
LHCI	light-harvesting complex of photosystem I.
LHCII	light-harvesting complex of photosystem II.
MBDoE	model-based design of experiments.

MPC Model Predictive Control.

NADPH nicotinamide adenine dinucleotide phosphate.

NER net energy ratio.

NLP nonlinear programming.

NLPSQP nonlinear programming sequential quadratic approach.

NPQ Non Photochemical Quenching.

OECD Organisation for Economic Co-operation and Development.

OMBRE online model-based redesign of experiments.

OPEX operation cost.

PAM pulse amplitude modulation.

PAR photosynthetically active fraction.

PBR photobioreactor.

Pc plastocyanine.

PQ plastoquinone.

PSI photosystem I.

PSII photosystem II.

PSU photosystem unit.

qE energy-dependent quenching component.

qI inhibition-dependent quenching component.

qT state-transition quenching component.

ROS reactive oxygen species.

SQP sequential quadratic programming.

TEF total electron flow.

WD Water Demand.

Nomenclature

Greek letters

α statistical degree of significance.

α_F activity level of fast energy dependent quenching.

α_S activity level of slow energy dependent quenching.

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

α_a thermal diffusivity of air.

α_{vs} power law exponent in wind velocity scaling.

β constant term for t^{up} threshold definition.

$\beta_{reg,SS}$ stationary value of cyclic flow deactivation efficiency.

γ constant term for N_{sw} threshold definition.

ξ_F time constant of the fast NPQ activation-relaxation mechanism.

α_{reg} cyclic flow deactivation efficiency.

ξ_{reg} activation/deactivation constant for cylcic/acidification rate regulation.

ξ_S time constant of the slow NPQ activation-relaxation mechanism.

δ constant term for N_{sp} threshold definition.

ΔT_{ml} logarithmic mean temperature difference between hot and cold side of the heat exchanger.

ΔT_{rw} temperature difference between hot and cold side of the heat exchanger.

ε_a emissivity of air.

ε_w emissivity of water.

σ_{H_s} std. deviation of solar irradiance measurements.

σ_{SB} Stefan-Boltzmann constant.

σ_{T_a} std. deviation of air temperature measurements.

η_D rate of basal thermal decay relative to the rate of fluorescence.

η_H fraction of PAR in solar light.

η_I rate of inhibition related quenching relative to the rate of fluorescence.

η_P rate of photoproduction relative to the rate of fluorescence.

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

$\bar{\eta}_{qE}^{C}$ maximum rate of interaction energy dependent quenching relative to the rate of fluorescence.

$\bar{\eta}_{qE}^{F}$ maximum rate of fast energy dependent quenching relative to the rate of fluorescence.

$\bar{\eta}_{qE}^{S}$ maximum rate of slow energy dependent quenching relative to the rate of fluorescence.

η_{reg} efficiency parameter for cyclic/acidification rate regulation.

φ_i i-th element of the design vector.

γ_{LK} correlation term in $\tau_{D,c}$ definition.

γ_x scaling concetration index.

λ_a thermal conductivity of air.

λ_r maximum specific respiration coefficient.

$\lambda_{r,max}$ maximum value of λ_r.

μ_m maximum specific growth rate.

$\mu_{m,max}$ maximum value of μ_m.

ν_a kinematic viscosity of air.

ω hour angle.

ω_s absolute value of the hour angle at sunrise/sunset.

ϕ_f fluorescence quantum yield.

ϕ_T temperature dependent function in growth model.

ψ $\mathbf{V}_\theta/\mathbf{H}_\theta$ measurement function for design optimization.

ρ_s soil density.

ρ_w pond medium density.

σ standard deviation on measured variables.

σ_b extinction coefficient.

σ_{θ_i} standard deviation on the i-th estimated model parameter.

σ_{PSI} effective cross section of photosystem I.

σ_{PSII} effective cross section of photosystem II.

σ_{tc} total cross section.

τ experiment/test/simulation/optimization duration.

$\tau_{D,c}$ $H_{D,c}/H_0$ ratio.

$\tau_{d,c}$ $H_{d,c}/H_0$ ratio.

τ_n initial time for the n-th time horizon.

τ_{PSII} turnover time.

θ_i i-th model parameter.

$\hat{\theta}_i$ i-th estimated model parameter.

θ_z zenith angle.

ε_{rnd} randomly generated error component for T_a measurements.

ε_{T_a} measurement error of a T_a sample.

ζ conversion factor.

Matrices and arrays [dimension]

$\mathbf{C}$ set of constraint function $[N_C]$.

$\mathbf{G}$ set of active constraints $[N_C]$.

$\mathbf{g}$ dynamical system.

$\mathbf{H}_\theta$ dynamic information matrix $[N_\theta \times N_\theta]$.

$\tilde{\mathbf{H}}_\theta$ information matrix associated to a single updating interval $[N_\theta \times N_\theta]$.

$\mathbf{H}_\theta^0$ preliminary information matrix $[N_\theta \times N_\theta]$.

$\mathbf{H}_{\theta|j}$ information matrix for the j-th redesign $[N_\theta \times N_\theta]$.

$\mathbf{L}$ constant information matrix for the redesign $[N_\theta \times N_\theta]$.

$\mathbf{M}_k$ information matrix for the k-th sample $[N_\theta \times N_\theta]$.

$\boldsymbol{\varphi}$ design vector $[n_\varphi]$.

$\boldsymbol{\varphi}_j$ design vector before the j-th update $[n_{\varphi,j}]$.

$\boldsymbol{\varphi}^l$ lower threshold for the design vector $[n_\varphi]$.

$\boldsymbol{\varphi}^u$ upper threshold for the design vector $[n_\varphi]$.

$\boldsymbol{\varphi}^{opt}$ optimal design vector $[n_\varphi]$.

$\mathbf{q}$ inflow/outflow rate control vector [2].

$\boldsymbol{\Sigma}_\theta$ prior information matrix $[N_\theta \times N_\theta]$.

$\boldsymbol{\Sigma}_y$ variance-covariance matrix of measurements $[N_y \times N_y]$.

$\boldsymbol{\Theta}$ parameter domain $[N_\theta]$.

$\boldsymbol{\theta}$ vector of true model parameters $[N_\theta]$.

$\hat{\boldsymbol{\theta}}$ vector of estimated model parameters $[N_\theta]$.

$\boldsymbol{\theta}^0$ vector of initial guess of model parameters $[N_\theta]$.

$\mathbf{t}^{sp}$ vector of sampling points $[N_{sp}]$.

$\mathbf{u}$ vector of manipulated inputs $[N_u]$.

$\mathbf{V}_\theta$ variance-covariance matrix of model parameters $[N_\theta \times N_\theta]$.

$\mathbf{w}$ vector of time-invariant controls $[N_w]$.

$\boldsymbol{\xi}$ state vector.

$\mathbf{x}$ vector of state variables $[N_x]$.

$\hat{\mathbf{x}_k}$ vector of states estimation at the k-th instant $[N_x]$.

$\mathbf{x}^0$ vector of initial states $[N_x]$.

$\dot{\mathbf{x}}$ vector of derivatives on state variables $[N_x]$.

$\mathbf{y}$ measurements vector $[N_y]$.

$\hat{\mathbf{y}}$ vector of estimated responses $[N_y]$.

$\mathbf{y}^0$ vector of initial conditions $[N_y]$.

Symbols

A_g glucose content in the meal.

A_{he} heat exchanger area.

A_{LK} pressure-corrected air mass in $\tau_{D,c}$ definition.

A fraction of PSII units in open state .

A_{st} decanter area.

B fraction of PSII units in closed state .

CC cloudiness.

CCF capital change factor.

C_{death} 3σ case constraint on T_p.

C_g blood glucose concentration.

$C_{g,b}$ basal blood glucose concentration.

C_{gsc} subcutaneous glucose concentration.

$C_{gsc,b}$ basal subcutaneous glucose concentration.

$C_{gsc,tot}$ total subcutaneous glucose concentration.

C fraction of PSII units in inhibited state .

c_{p_s} soil heat capacity.

c_{p_w} water heat capacity.

C_{rw} cost of refrigerated water.

C_v process cost variable.

D_{death} 3σ case additional constraint.

l_p open pond depth.

$l_{p,\max}$ upper bound of l_p.

$l_{p,\min}$ lower bound of l_p.

$D_{meal}(t)$ dynamic meal disturbance.

$D_{w,a}$ mass diffusion coefficient of water vapor in air.

$E(\cdot)$ expected value.

f differential and algebraic system implicit function.

f_a theoretical photosynthetic efficiency.

F_c corrective factor.

f_{comp} compensation function.

F_l fluorescence flux.

G specific growth rate function.

G_d constraint for hypoglycaemic condition.

h measurement selection function.

H_c clear-sky total irradiance.

h_{conv} heat transfer coefficient.

$H_{D,c}$ direct component of the clear-sky total irradiance.

$H_{d,c}$ diffuse component of the clear-sky total irradiance.

H_s solar irradiance.

h_{st} decanter depth.

H_0 solar irradiance reaching the external surface of atmosphere.

I_c variation of plasmatic insulin concentration with respect to the basal value.

$I_{c,b}$ basal plasmatic insulin concentration.

IC_{cn} installation cost associated to the centrifuge.

IC_{he} installation cost associated to the heat exchanger units.

IC_{st} installation cost associated to the settler.

IC_{uo} installation cost associated to a generic unit operation.

I_{qE} Irradiance level at which half of the maximal qE activity is realized.

I solar irradiance.

I_{sc} solar constant term.

k_{cyc} ferrodoxin rate constant.

k_{cyt} cytochrome rate constant .

K_d Erbs correlation term.

k_d damage rate constant .

k_{ε} reduction coefficient for the weight term w_j.

K_I half-saturation parameter in the growth model.

K_m mass transfer coefficient.

k_{pq} plastoquinone rate constant .

k_r repair rate constant .

k_s soil conductivity.

k_T Erbs correlation term.

kt^* clear-sky index.

k_w incremental coefficient for the weight term w_j.

$K_{I,max}$ maximum value of K_I.

L_{ch} characteristic pond length.

$l_{s_{ref}}$ reference depth.

L_w water latent heat.

MC_{cn} manufacturing cost associated to the centrifuge.

MC_{he} manufacturing cost associated to the heat exchanger units.

MC_{uo} manufacturing cost associated to a generic unit operation.

m_e evaporation mass flux.

$M\&S$ Marshall & Swift cost index.

M_w molecular weight of water.

N_C number of constraints.

N_{chl} chlorophyll specific number of photosynthetic units.

N_D number of constraints on estimated outputs.

N_{day} Julian day of the year.

$n_{day,w}$ number of days considered for a generic simulation.

N_E number of constraints on unmeasured state variables.

N_{he} number of heat exchangers.

n_{φ} number of design variables.

n_{φ_i} number of design variables for the i-th redesign.

N_{pond} number of ponds.

n_{qE} Hill parameter related to the shape of sigmoid function describing NPQ activity.

N_{sc} number of scenarios.

N_{sp} number of samples.

N_{sw} number of switching points for piecewise constant manipulated variables.

N_{sw_j} number of switching points used for the j-th experiment.

Nu Nusselt dimensionless number.

N_u number of manipulated variables.

n^{up} number of redesigns.

N_w number of time invariant controls.

N_{week} number of weeks of plant operation.

N_x number of state variables.

N_y number of measured variables.

N_{θ} number of measured variables.

P_a saturated vapor pressure at T_a.

P_{net} net productivity.

Pr Prandtl dimensionless number.

P_w saturated vapor pressure at T_p.

$q^{buf,max}$ q^{he} value for the *With Buffer* case studies.

Q_{cond} conductive flow with the ground at the pond bottom.

Q_{conv} convective flow at the pond surface.

Q_{ev} evaporation flow.

Q_{he} heat exchanger duty.

q^{he} volumetric flowrate in the heat exchanger.

q_i i-th component of $\mathbf{q}$ vector of inputs.

q^{in} inlet flowrate.

$q^{in,max}$ maximum value of q^{in} during the optimization window.

Q_i heat flow due to the water inflow.

$q_{\max}$ upper bound of $\mathbf{q}$ vector components.

q^{out} outlet flowrate.

$Q_{ra,a}$ radiation flow form the air to the pond.

Q_r heat flow associated with rain.

$Q_{ra,p}$ radiation flow form the pond surface.

$Q_{ra,s}$ total (direct + diffuse) solar irradiance.

q^{rw} volumetric flowrate of refrigerated water.

q^{st} settler outlet flowrate.

R specific respiration function.

n_{cytc6}/n_{PSII} cytc6/PSII molar ratio.

$x_{cytc6,1}(t = 0)$ initial value of reduced Cytc6.

n_{fd}/n_{PSII} fd/PSII molar ratio.

Re Reynolds dimensionless number.

R_g universal ideal gas constant.

r_g growth rate used in IDRO case study.

RH relative air humdity.

R_{ut} tissue rate utilization.

S open pond surface.

Sc Schmidt dimensionless number.

S_f scaling factor for fluorescence model.

Sh Sherwood dimensionless number.

$\sigma_{PSII}/\sigma_{PSI}$ antenna size ratio.

s_{ij} ij-th element of the inverse matrix of measurements errors.

t time.

T_a air temperature.

$T_{a,avg}$ weekly averaged air temperature.

TC_{cn} total cost associated to the centrifuge unit.

TC_{he} total cost associated to the heat exchanger units.

TC_{st} total cost associated to the settler unit.

TC_{uo} total cost associated to a generic unit operation.

$t(\cdot)$ t-value distribution.

t_i i-th t-value.

T_f total duration of the optimization.

T_h time horizon length.

T_{in} water inflow temperature.

T_{LK} empirical term in $\tau_{D,c}$ definition.

t_{op} plant operation duration.

T_p open pond temperature.

T_{max} temperature above which growth is zero.

T_{min} temperature below which growth is zero.

T_{opt} temperature at which $\mu_m = \mu_{m,max}$.

T_{p_σ} artificial T_p variable built with wrong weather data.

T_r reference computational time.

$t^{ref}_{1-\alpha}$ reference t-value at (1-α)% confidence.

T_s soil temperature.

$t_{sp,i}$ i-th sampling time.

$T_{s_{ref}}$ soil temperature at the reference depth $l_{s_{ref}}$.

t^{up} updating time point.

$t^{up,j}$ updating time point before the j-th redesign.

u generic manipulated input.

U_{he} heat exchanger global heat transfer coefficent.

V open pond volume.

V_G glucose distribution volume.

V_I insulin distribution volume.

v_{ii} ii-th element of the parametric variance-covariance matrix.

V_{max} maximum daily volume extracted by the pond.

v_r rainwater flow.

V_τ final open pond volume.

v_w wind velocity.

V_0 initial open pond volume.

v_0 wind velocity at height $z_{m,0}$.

w_j weight term for IDRO objective function calculation.

x generic state variable.

x_b biomass concentration.

$x_{b,opt}$ optimal biomass concentration.

X_I variation of insulin concentration in the sub-compartment with respect to the basal value.

y generic measured output.

z depth coordinate.

z_m wind velocity height.

$z_{m,0}$ wind sensor height.

List of Figures

List of Tables

Chapter 1

Introduction

Microalgae processes have emerged during the last decade as one of the most promising new technologies for providing innovative molecules for the cosmetic and pharmaceutical industry, and as a source of proteins for animal and human nutrition (Mata et al. (2010); Skjånes et al. (2013)). At a larger time horizon, microalgae are expected to contribute for fossil carbon replacement with renewable carbon, especially for supplying green chemistry and liquid biofuel in the transport sector (Foley et al. (2011b)). The great interest in this technology is not only related to the substantial higher productivity compared to terrestrial plants (Chisti (2007)), but also to the possibility of coupling the microalgal production process to industrial CO_2 mitigation and wastewater treatments to finally recycle carbon, nitrogen and phosphorus. Nevertheless, much research is still needed in order to make this potential new energy source a real feasible technology, since all the existing techno-economic assessments (Wigmosta et al., 2011; Yang et al., 2011; Moody et al., 2014) were limited by uncertainties regarding the biomass productivity that can actually be reached at full-scale.

In fact, while few assessments based productivity assumptions on model predictions, most of the existing studies used experimental data collected either indoors or at certain time of the year to extrapolate outdoor algal yields over an entire year of cultivation, then leading to significant differences in assessment conclusions regarding the sustainability of full-scale algal cultivation. A better understanding of the multiple time-scale underlying biophysical and chemical processes is therefore necessary: firstly, an accurate description of solar irradiance, temperature and nutrients mutual interaction and impact on microalgae growth is required. Secondly, a complex analysis of the cultivation system design and management is needed, by taking into account macroscale phenomena, such as macro-mixing for optimal nutrients uptake, light distribution optimization, external effects on culture medium (evaporation, contamination,..).

The work presented in this Thesis aims at investigating microalge growth behavior at two different scales: firstly, the 'first-principles' state model of Bernardi et al. (2016), describing microalge photosynthetic mechanisms, was extended by implementing the electron transport chain process. Then, the productivity optimization of an open pond system through meteorological data was investigated at different seasons, in order to both understand the system behavior and propose a reduced set of practical guidelines for future operation. Finally, different approaches to cope with wrong system control manipulation caused by inaccurate weather forecasts were proposed and discussed. The main objective of this Chapter is to present the aim of the Thesis research activity. First, a general overview of the energy situation is presented to discuss of the state of the art in microalgae-based processes; then, a focus on microalgae as renewable energy sources is discussed together with the main production process alternatives. Next, a focus on the cultivation phase is presented, with a brief comment on the advantages and drawbacks associated to current cultivation systems. Finally, a literature review of microalgae dynamics modelling is assessed. The motivation of the work and a scheme of the Thesis structure will conclude this Chapter.

1.1 Towards sustainable energy and chemicals production

Energy usage has been considered one of the most important and ongoing issues of the modern time: the continuously growing enhanced living standards of developed countries (Organisation for Economic Co-operation and Development (OECD)) and the rapid population growth (Roser and Ortiz-Ospina (2017)), especially in non-OECD realities (such as India, China, sub-Saharian Africa and Middle East countries), result in a rapid rise in global energy demand. In order to fulfill the global energy requirements (13699 Mtoe in 2014, according to IEA (2016)), consumption of fossil fuels has increased; in particular the world demand of crude oil, coal and natural gas accounted for 81.1% of total energy demand in 2014 (31.3%, 28.6% and 21.2%, respectively, see Fig.1.1). Table 1.1 shows that global energy demand is predicted to grow at two different rates: at a lower average annual change (0.6%) for already developed countries and at a rapid average speed in emerging Asiatic and African countries (with a maximum peak of 3.2% for India). This dramatic trend towards an ever increasing energy consumption triggers different questions about our near future:

Table 1.1: *Actual and predicted annual energy consumption (in quadrillion BTU) world-wide by geographic area. Data collected from EIA (2016) website.*

	Region	yr 2016	yr 2025	yr 2040	Avg. Annual Percent Change (yr 2012 - yr 2040)
	OECD	246.7	260.6	282.1	0.6%
Non-OECD	Asia	201.1	246.4	279.9	2.2% (India + 3.2%)
	Africa	23.3	30.0	44.0	2.6%
	Middle East	35.2	45.4	61.8	2.4%
	Central and South America	31.0	36.7	47.3	1.5%

- **Finiteness of fossil resources.** Fossil fuel reserves are likely to be rapidly depleted; Shafiee and Topal (2009) proposed a modified version of Klass (1998) model describing fossil fuels consumption and concluded that fossil fuel reserves depletion for oil, coal and natural gas is, respectively, 35, 107 and 37 years, starting from 2005. In other words, according to this prediction, coal reserves will be available up to 2112, and will be the only fossil fuel remaining after 2042. More pessimistic predictions for coal depletion were given by Mohr et al. (2015): in this work different scenarios at different risk probability were simulated for four countries (China, USA, Canada and Australia) with a final estimation of coal production peak before 2025 due to China industrial production. Focusing on crude oil production, Nashawi et al. (2009) stated that, on the basis of 2005 world crude oil production and recovery techniques, the world oil reserves will be depleted at an annual rate of 2.1%. This is due to the fact that oil demand has increased in the past few years because of the rapid growth in the transportation sector, in addition to the absence of viable economic alternatives for fossil fuel. Even though different forecasting assessments on fossil fuels depletion are given in literature, many experts agree that fossil resources supply cannot satisfy the growing demand of energy by itself.

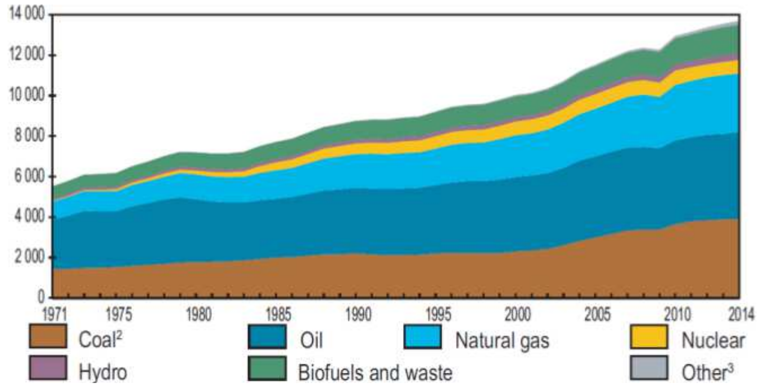


Figure 1.1: *World primary energy assessment (IEA (2016))*

- **Climate Changes.**

In order to maintain the thermodynamic balance of the Earth at an equilibrium temperature, the amount of energy absorbed as solar radiation should be kept equal to the amount of energy emitted back into space at infrared wavelenghts (Guerrero-Lemus and Martinez-Duart (2013)). Nevertheless, the so-called greenhouse gases (GHG, mainly CO_2, CH_4 and NO_x, accounting for 73.5%, 19.0% and 5.9% of the total GHG emissions, respectively, see EIA (2016)) in the atmosphere absorb and re-emit infrared radiation, hence warming up the lower atmosphere and the Earth's surface with negative consequences for life and environmental protection (Pann, 2011). The increasing use of fossil fuels (especially due to industry (29.1%), building (18.3%) and transport sector (14.5%) demand, see Fig.1.2) has been the main responsible of the continuous increase of GHG emissions (51840 MtCO_2eq in 2016). A lot of policies have been recently developed in order to cope with climate change issues; one of the most recent examples in Europe is The Paris Agreement, which entered into force on 4 November 2016 to set a major step forward in the fight against global warming. The main target consists of holding the increase in the global average temperature to well below 2 °C above pre-industrial levels and to pursue efforts to limit the temperature increase to 1.5 °C above pre-industrial levels. Yet, meeting such ambitious climate goals will be extremely challenging and would require a step change in the pace of decarbonisation and efficiency. Despite Paris Agreement theoretical targets, IEA (2016) reports that implementing current international pledges will only slow down the projected rise in energy-related carbon emissions from an average of 650 million tonnes per year since 2000 to around 150 million tonnes per year in 2040. A lot of technical assessment were done in order to test the practical feasibility of Paris Agreement topics: as an example, Meinshausen et al. (2009) made a comprehensive probabilistic analysis aimed at quantifying GHG emission budgets for the 2000-2050 period that would limit warming throughout the twenty-first century to below 2 °C, based on a combination of published distributions of climate system properties and observational constraints. According to their results, the probability of exceeding 2 °C would be within 53-87% if global GHG emissions are still more than 25% above 2000 levels in 2020. Unfortunately, despite the pledges made for the Paris Agreement on climate change, the era of fossil fuels appears far from over (natural gas continues to expand its role, while the shares of coal and oil fall back) and underscores the challenge of reaching more ambitious climate goals. The path to 2 °C is tough (IEA, 2016), and it can be achieved only if policies to accelerate further low carbon technologies and energy efficiency are put in place across all sectors.

- **Risk to living creatures on the Earth.** Climate change has produced numerous shifts in the distributions and abundances of animal species over the past 30 years. Thomas et al. (2004) assessed the extinction risk for sample regions covering 20% of the Earth's surface and predict, on the basis of mid-range climate-warming scenarios for 2050, that 15-37% of species will be 'committed to extinction'. These estimates show the importance of rapid implementation of technologies to decrease greenhouse gas emissions and strategies for carbon sequestration.

In order to keep the Earth safe and to counter the environmental/economical threats, sustainable and pollutant-free technologies have been introduced, known as renewable sources. Renewable energies (i.e. solar energy, wind energy, geothermal energy, biomass energy, biofuels,..) can reduce the planet thermal imbalance, limiting greenhouse gases increase in the atmosphere. Currently, the above cited renewable energy sources supply about 13.8% of total world energy demand (IEA, 2016); moreover, new emerging renewable technologies are under investigation, such as marine energy, concentrated solar photovoltaic (CSP) and enhanced geothermal energy (EGE) (Hussain et al., 2017). Renewables had an impressive development worldwide, with highest growth of solar photovoltaic (42% annual growth over the last decade) and wind (27% annual growth, see IEA (2016)). In related terms, the most significant growth was made by the renewable energy use in transport

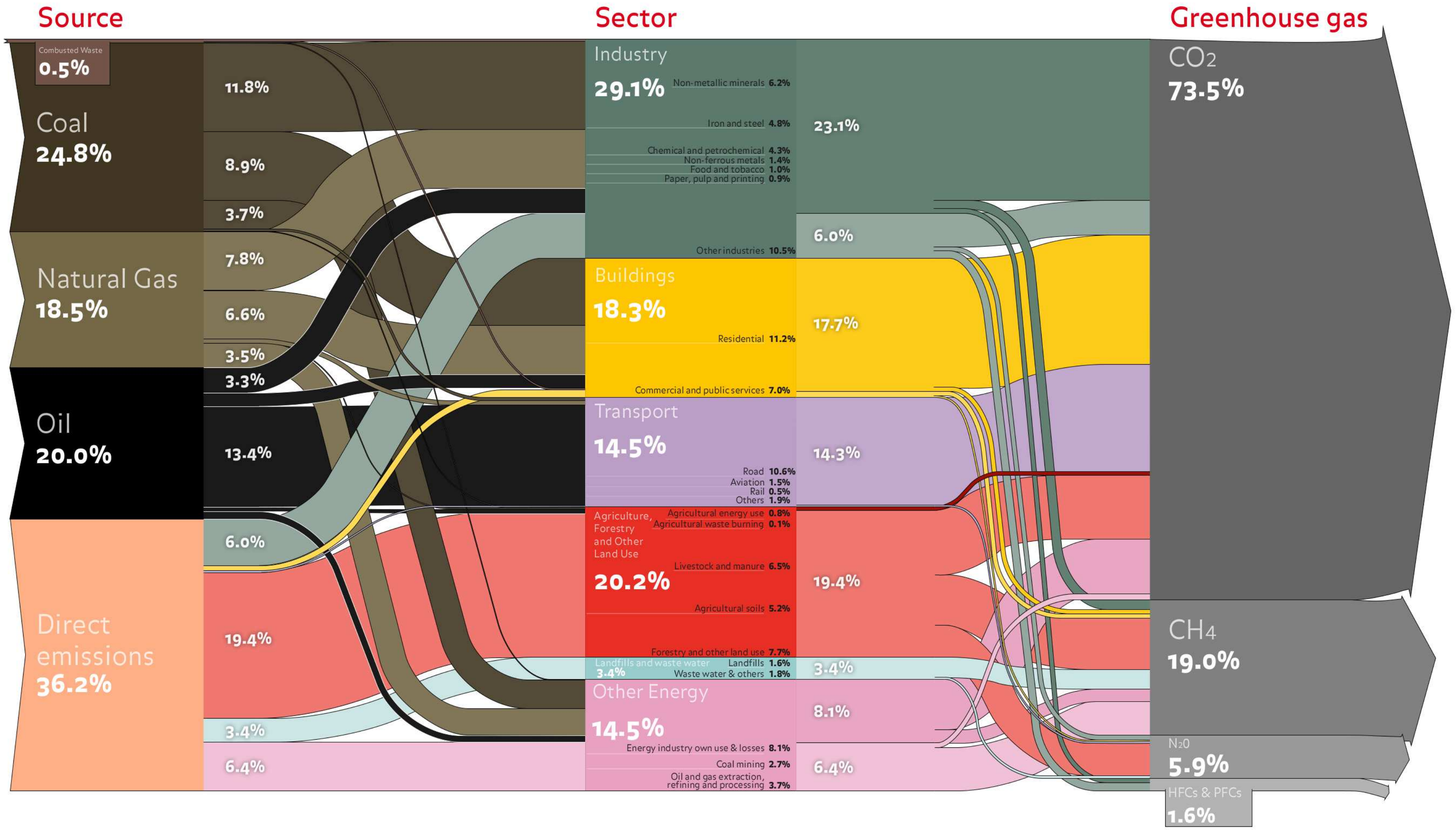


Figure 1.2: *Greenhouse gases flowchart (EIA (2016)).*

section (+223% between 2005 and 2012); the use of biomass-based biofuels has continuously increased to 106 billion liters worldwide in 2012 (82.6 of ethanol and 23.6 of biodiesel). The main driver in this field is related to the fact that the transport sector is almost totally dependent on crude oil (more than 97% of total supplies for this specific sector) and new technologies must be rapidly developed to cope with crude oil depletion. After the initial great interest on first/second generation of biofuels, based on traditional crops or ligno-cellulosic materials, a new third generation of biofuels based on microalgae is now under investigation, given its potential as a new renewable energy and chemical source. The current potential of microalge-based biofuels is mainly based on its high theoretical productivity, if compared to terrestrial plants. Microalgae, in fact, can potentially convert 13% of total solar light into chemical energy through photosynthesis (Smil, 2008). Nevertheless, microalgae cultivation in artificial systems dramatically reduces the actual light-to-energy conversion to 1%-3%, depending on the chosen cultivation system (Norsker et al., 2011). Further process understanding is therefore required to quantify the key variables affecting growth and metabolic mechanisms, in order to improve the current process efficiency. Recent scientific breakthroughs should be mentioned about fourth generation biofuels and the so called ‘synthetic biology’, enabling direct conversion of solar energy to fuels from sun light, water and CO_2. Nonetheless, further maturity of synthetic biology as a technology is required in order to make biology fully engineerable and production systems economically viable (Aro, 2016). For this reason, the work presented in this Thesis deals with third generation biofuels. In fact, microalgae processing currently represents one of the most promising technologies for both sustainable production of a wide range of high values products and biofuels in the transport sector.

1.2 Renewable energy sources: microalgae

Microalgae are microscopic, unicellular/multicellular prokaryotic/eukaryotic photosynthetic organisms that can produce biomass and oxygen by using sunlight as energy source, CO_2 as carbon source and inorganic salts as nutrients. They can guarantee higher growth and photosynthetic rates with respect to other plants, for their all-year production capability (Kiran et al., 2014). Furthermore, they can grow under extreme environmental conditions, low nutritional and water requirements and without herbicides or pesticides. Among the different cultivation technologies for bulk production of microalgae biomass, phototrophic cultivation is currently one of the most investigated ones for large scale microalgae biomass production; it means that the key fundamental process responsible of microalgae-based chemical energy generation is photosynthesis.

The photosynthetic process (Eberhard et al. (2008); Antal et al. (2013)) is commonly split into two main phases: the first one is the light-induced electron transport chain for adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH) generation. In this phase oxygen is produced through water splitting as a side product

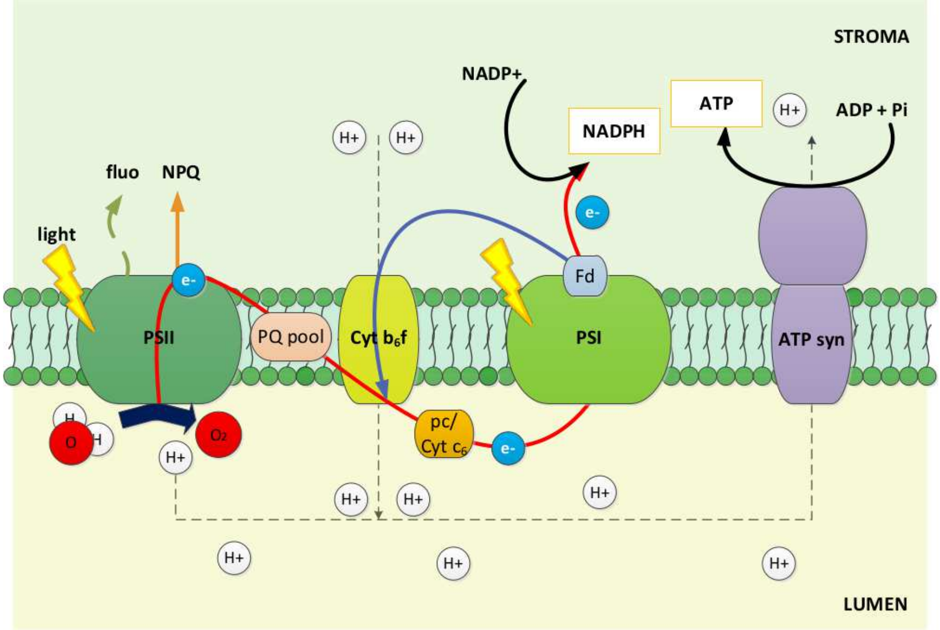


Figure 1.3: *Photosynthetic dynamics representation*

(see Fig.1.3). ATP and NADPH energy carriers are used in a second stage, independent of light, during which the energy stored in the ATP and NADPH molecules is used to fix CO_2 and product sugars or other molecules. In microalgae the light-dependent phase occurs in the thylakoids membrane, a phospholipidic membrane that separates two zones called lumen (the inner region) and stroma (the outer one). Two main complex proteins (LHCII and LHCI), binding the bulk of chlorophyll and carotenoids of the entire cell, are responsible of light capturing, harvesting and excess energy dissipation in case of excess of solar irradiation. The light energy captured by LHCII and LHCI is transported, respectively, to the reaction centres (photosystem II (PSII) and photosystem I (PSI)) via a coordinated network of pigments. The energy absorbed by LHCII for PSII reaction centre activation is used to drive the water splitting reaction. This reaction turns water into electrons, protons, and oxygen. The electrons are transported along the membrane by a series of electron carriers (plastoquinone (PQ), cytochrome b6f complex (Cytb6f), plastocyanine (Pc), ferrodoxin (Fd) and PSI), following the so-called linear electron flow (LEF). The final acceptor is generally the $NADP^+$ molecule, which is converted in NADPH, an energetic molecule used in the dark phase of photosynthesis. At the same time, a proton flux is built between stroma and lumen thanks to PQ oxidation, with a consequent acidification of the lumen. The proton gradient drives the ATP production in the ATP synthase. An alternative electron path is the cyclic electron flow (CEF), driven by PSI. The cyclic reaction is similar to the linear reaction but produces only ATP and no NADPH is generated. Furthermore, once the electrons leave PSI they are passed down through the electron carriers and returns to PSI through a different path (blue arrow in Fig.1.3). Photosynthetic dark reactions occurs in the stroma region of thylakoids through a 3-step process referred as Calvin-Benson cycle, which uses ATP and NADPH generated by the light reactions to fix CO_2.

The growth rate and the maximum biomass production of microalgae cultures does not depend uniquely on light. In fact, microalge culture is also affected by abiotic (temperature, pH, salinity, oxygen, nutrients, and toxic chemicals), biotic (pathogens and competition by other algae), and operational (mixing, depth control, harvest frequency, ...) factors (Mata et al., 2010).

Depending on the species (Medipally et al., 2015), microalgae are capable to produce varying amounts of lipid, polyunsaturated fatty acid, natural dye, carotenoid, antioxidant, enzyme polymer, peptide, toxin and sterols (Moreno-Garcia et al., 2017), with potential applications in different industrial sectors (biofuels, cosmetics, pharmaceuticals, nutrition and food additives, aquaculture and pollution prevention). Valuable co-products, such as proteins and pigments, could be obtained from microalgae through biomass composition modulation by modifying the nutritional requirements/growth conditions/process technologies (Bona et al., 2014) according to the desired final use (see Patel et al. (2017)). As you can see in Fig.1.4, culture harvesting (and dewatering) follows the cultivation phase. The algal biomass is usually harvested through centrifugation, flotation or filtration techniques. At this step, several process strategies can be adopted, depending on the desired final use. For example, ethanol and bioplastics can be obtained with direct fermentation; if the user is interested in biohydrogen production, algal biomass can be gasified and converted through a Fischer-Tropsh process. Finally, the biomass could be used as a direct energy source through direct combustion. If we want to produce biofuels (through lipids transesterification into methyl esters) or high value compounds

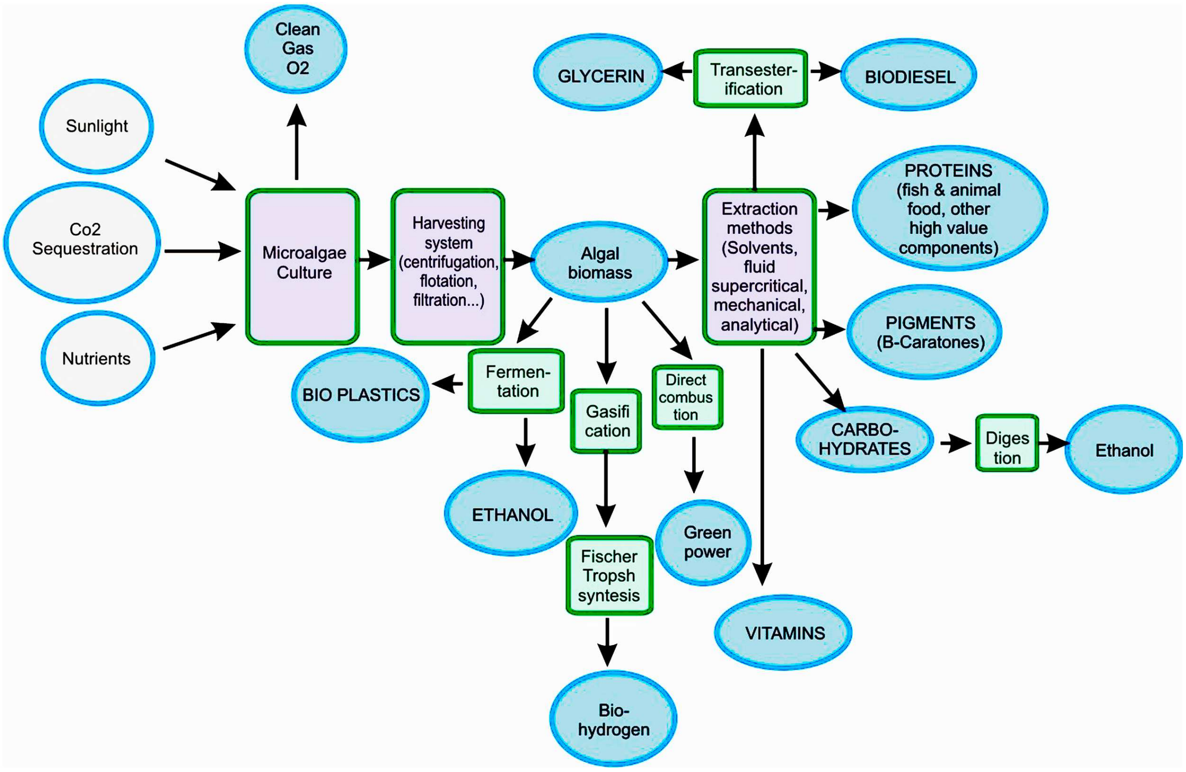


Figure 1.4: *Microalgae-based processes flowchart (Patel et al. (2017)).*

(such as proteins, pigments, carbohydrates), an additional phase of extraction is required; the three most common alternatives are the mechanical expeller, the extraction with solvent (usually hexane) and supercritical CO_2 fluid extraction (Krichnavaruk et al., 2008). Following a path rather than another requires deep knowledge of the right cultivation and pretreatment conditions, since high biomass productivity values is often associated to low lipid/carbohydrates content. It is therefore necessary to gain better understanding of the trade-offs between lipid and co-products accumulation and productivity (Araya et al., 2014).

Since in this Thesis work we will almost exclusively focus on the cultivation phase of microalge production process, a brief description of the most common cultivation technologies is given in the following paragraph.

1.2.1 Microalgae cultivation technologies

If we considered phototrophic cultivation, the most commonly used systems are the so-called open ponds or the enclosed photobioreactors (PBRs).

- **Open Ponds**. Open ponds are the oldest (practiced since the '50s, as stated in Borowitzka (1999)) and simplest cultivation systems used for large scale microalgae production (covering about 98% of commercial algae production in 2015). The open pond systems differ in terms of size, shape, building materials, mixing equipment, and inclination with respect to the ground surface (Borowitzka, 2005). Raceway ponds can be stirred by a paddle wheel (the most common option, see Jiménez et al. (2003)) or used as extensive shallow unmixed systems. In any case, open pond systems are the cheapest technology for large scale cultivation of microalgae compared to close PBRs. Furthermore, cultivating in open ponds does not compete with agricultural crops for land (Chisti, 2007), does not require frequent maintenance and cleaning and consume relatively low energy (Ridolfi et al., 2009). However, there are some drawbacks in terms of final productivity, such as the strong dependence on highly fluctuating local conditions (temperature, light intensity, ...) and high probability of contamination from the air and ground (Belay, 1997).

- **Enclosed Photobioreactors (PBRs).** PBRs are generally available in the form of glass, plastic, or other transparent materials tubes, bags, or plates. Some common PBRs designs include annular, tubular, and flat-panel reactors, with large specific surfaces (Pulz, 2001). Photobioreactors present many advantages, such as high system efficiency given by major control over culture conditions (evaporation elimination, nutrients supply optimization), but their construction, operation, and maintenance cost is currently higher than open ponds. This problem still limits the cost-effective production of microalgae biomass on large scale.

Jorquera et al. (2010) recently made a comparative life cycle analysis (LCA) study in order to compare open ponds, tubular and flat photobioreactors in terms of net energy ratio

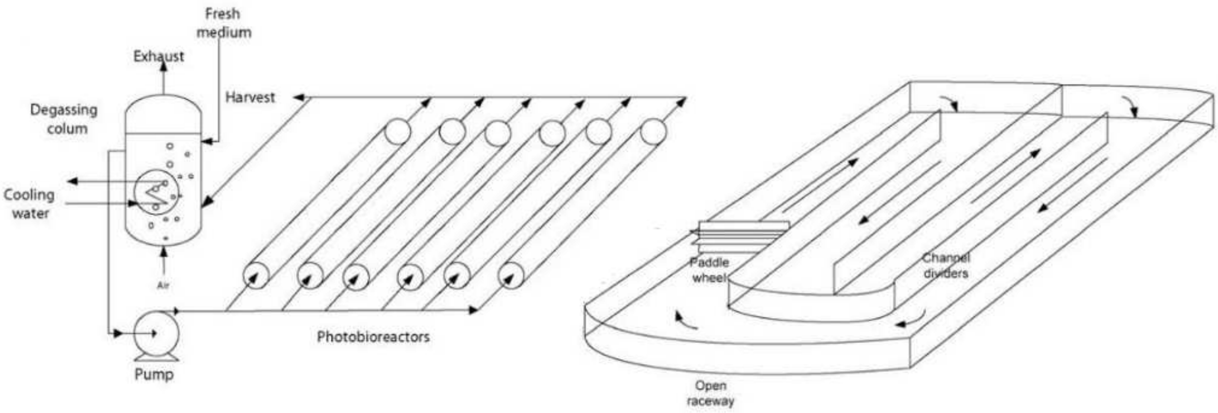


Figure 1.5: *Tubular photobioreactor and open pond schematisation. Figures are taken by Jorquera et al. (2010).*

(NER). NER is the ratio between the energy produced by the cultivation system and the energy required to operate it. The results showed that only raceways ponds and flat PBRs are economically sustainable, with NER > 1. A more pessimistic evaluation was made by Richardson et al. (2012); in this work a multi-year, Monte Carlo financial feasibility model was developed to estimate the costs of production and chance of economic success for large scale production with different cultivation systems. The financial feasibility analysis showed that the only way to achieve a 95% probability of economic success in the PBR system was to reduce CAPEX by 80% or more and OPEX by 90% or more. For the open pond system there were more options that could return a 95% or greater chance of economic success, for example, reducing CAPEX by 60% and OPEX by 90%.

The high uncertainty on microalgae cultivation practical feasibility at large scales requires further technological improvement, together with a deeper and an ever increasing understanding of the underlying mechanisms of microalgae growth. In this perspective, microalge growth modelling can be considered a valuable tool to support both system design and operation optimization, with consequent potential increase of cultivation profitability.

1.3 Microalge growth modelling: a multiscale problem

Microalgae cultivation processes exhibit high complexity due to the interaction of multiple phenomena that span multiple time scales, ranging from milliseconds to days, and from nanometers to the hundreds of meters in industrial cultivation plants, involving quantum phenomena, cell metabolism, turbulence and fluid flow behavior, light diffusion in a non-transparent medium, mass and energy transport phenomena, interactions with the equipment design and the external environment. Focusing on the photosynthetic mechanism, several sub-processes are responsible of the overall productivity: *photoproduction*, the collection of all processes from photons utilization to CO_2 fixation, that occurs in a fraction of a second; *photoinhibition*, the observed loss of photosynthetic production due to excess of light, which acts on time scales of minutes to hours; *photoregulation*, the set of mechanisms by which microalgae protect photosynthetically active components via the dissipation of excess energy as heat (Non Photochemical Quenching NPQ), that occurs within minutes; *photoacclimation*, the ability of microalgae to adjust their pigment content and composition under fluctuating light and nutrient conditions, acting on time scales of hours to days; and finally, the mechanisms of *nutrient uptake* that lasts within hours to days.

Several dynamic models of microalge growth have been developed in the last 60 years (especially for investigating oceanography, ecology and biotechnology scientific domains), from the simple description of microalgae growth rate through the hyperbolic function proposed by Baly (1935) up to the more complex representations. The first developed models were built to describe microalgae growth in their original environment, to which the specific strain adapted over billions years of evolution. The model proposed by Riley (1946), for example, was proposed to describe the population of Georges Bank by considering both the effect of light (through an exponential function representing the light decrease along the culture medium depth) and the impact of nutrient on growth. However, the artificial cultivation systems described in the previous paragraph must be considered as new sub-optimal systems to which microalge are not naturally adapted. Amongst the main differences with respect to cultivation in natural habitats we can highlight: (*i*) the high biomass concentrations usually required for large-scale production (microalge naturally grow at very low concentrations), (*ii*) the artificial administration of macronutrients to avoid limiting conditions and (*iii*) the cultivation temperature, often fixed at values (specially in PBRs) that guarantee maximum productivity for the specific strains considered.

Different modelling strategies have been adopted in the last 60 years (Bernard (2011)) to describe microalgae growth: from the first empirical models (see, as an example, the above mentioned model by Baly (1935), then extended by Vollenweider (1965)) based on empirical correlations, to the more sophisticated mechanistic representation of the underlying physical-biological processes through the so-called ‘first-principles’ models. If we focus our attention on this last class of models, two main categories can be identified: physiological and state models.

Physiological models aim to describe the dynamics of the most significant sub-processes involved in both photosynthesis and cell metabolism. The basic idea is to propose an approximation of the actual mechanisms of mass and energy storage/transport that are responsible of microalge growth. As an example, Baroukh et al. (2014) recently proposed a dynamic metabolic modelling framework that handles non-balanced growth conditions and accumulation of intracellular metabolites. The metabolic network was split into sub-networks describing spatially close reactions; then each sub-network was reduced to a limited set of macroscopic reactions with simple kinetics. Finally, an ordinary differential equation system was obtained to describe substrate consumption, biomass production, products excretion and accumulation of internal metabolites.

Physiological models are therefore extremely detailed and generally involve a large amount of variables and parameters, despite recent implementation of advanced techniques for model simplification. The identification procedure results extremely difficult, sometimes impossible, due to the high number of parameters used. For this reason, especially to describe the main key processes involving nutrient uptake and transport in the cell, simpler compartmental model structures were adopted in literature. Nutrient uptake models are usually based on Droop (1968) modelling approach (based on Dugdale (1967) nutrients uptake description and reformulated by Burmaster (1979)), initially used to represent the effect of vitamine B_{12} internal quota on the growth rate of phytoplankton and recently extended to the representation of macronutrients (nitrogen and phosphorus) effects on microalge growth rate at constant light. This model was adopted in several works; as an example, Mairet et al. (2011) used a modified Droop model in order to assess the hysteresis phenomenon in neutral lipid production: in practice, the lipids dynamics after nitrogen starvation is highly different from the dynamics after nitrogen recovery. Simplified compartmental models are highly useful to predict significant metabolic paths without considering unnecessary biological details for macroscale production. Furthermore, the simpler the model, the simpler the coupling with other models (light distribution, temperature,...) will be in terms of final identifiability.

In parallel to nutrient compartmental models, several kinetics models of photosynthesis have been proposed to describe the impact of light on growth. The effect of light on microalge is complex: in fact, as described at the beginning of this section, light irradiance is not only responsible of triggering CO_2 fixation through photosynthesis, but can also drive photoinhibitory phenomena. In particular, the excess of incoming light can be responsible for the production of reactive oxygen species (ROS) that damage D1 proteins, hence hindering the energetic efficiency of the photosynthetic process. Moreover, microalgae have evolved by developing a complex regulatory systems to protect themselves via heat dissipation of excess incoming energy. This is the above mentioned NPQ, whose dynamics is conventionally split into three components: the most rapid component (qE), depending on pH gradients between stroma and lumen and on xantophyll protein synthesis; the second component (qT), relaxing within minutes, associated to LHCII uncoupling from PSII ('state transition'); the slow third component (qI) related to photoinhibition-driven mechanisms (Müller et al. (2001)). The great complexity in describing microalgae behavior at variable light intensities has recently increased the interest on mechanistic models development, in particular on the so-called 'state-models'.

State models (Steele (1962), Peeters and Eilers (1978), Platt et al. (1980), Han (2002)) are based on the concept of PSU unit, which consists of the antenna complex and the associated reaction center, together with the stoichiometric apparatus that is activated by a specific amount of light to produce a given amount of NADPH. The name 'state model' was coined to reflect that PSUs can be in different states of excitation (open, closed and light-damaged). Wu and Merchuck (2001) model, as an example, is based on Peeters and Eilers (1978) assumption that only the closed centers can be damaged by light in excess; then, the damaged PSUs centers can be directly recovered to open state through a first-order kinetics.

Conversely, Nikolau et al. (2015) changed the model structure (as proposed by Han (2002)) assuming that the damaged PSUs centers are first recovered to the closed state through a first-order kinetics and then to the open state with a fast reaction. Another approach was tested by Bernardi et al. (2014); in this work the model proposed by Camacho-Rubio et al. (2003), based on the assumption that photoinhibition affects both open and closed PSUs, was simplified and modified through the implementation of photoregulation mechanisms. An interesting contribution was proposed by Papadakis et al. (2012), who described the cyclic transport through the introduction of a fourth state.

In the past decade several models have also been proposed to include another key regulatory mechanism, photoacclimation. This mechanism accounts for the response of microalgal pigment/protein density to both light intensity and limiting nutrients and acts on a timescale of days to weeks. In practice, it consists on the cell capacity to reduce its pigment content in case of exposition at high light irradiances. Geider et al. (1998) were the first to propose a simple photoacclimation representation introducing chlorophyll as model variable (in addition to microalgal carbon and nitrogen) and integrating the known response of photosynthesis to both light and nitrogen status in the cell. An important contribution was given by Flynn (1991) with a growth model considering the light effects on growth, photoacclimation phenomena and also the growth dependence on temperature. Other models have been proposed for photoacclimation description (Zonneveld (1998) and Pahlow (2005)) by implementing more complex details on the fundamental mechanism. Recently, Gárcia-Camacho et al. (2012) proposed a novel model formulation taking into account the dynamics of photoacclimation, the effect of non-photochemical quenching as a response to high irradiation, as well as other aspects such as dark respiration. The model recently proposed by Ebenhöh et al. (2014) for *Chlamydomonas* recapitulated the basic fluorescence features of short-term light acclimation (state transitions); furthermore, a photosynthetic electron transport chain representation was coupled to a heuristic description of CEF to simulate qE triggering.

Typically, temperatures effects on microalgae growth have been studied following two main approaches (Béchet et al. (2013)): the first one considers light and temperature as two independent variables acting on growth (uncoupled approach), whereas the second one aims at representing the interdependency between the two variables (coupled approach); this last approach has been used by Duarte (1995), who proposed a state model whose dynamics were described through Arrhenius-like kinetics. Although coupled models theoretically represent temperature effects on growth with a detailed description of the underlying mechanisms, uncoupled models are usually preferred due to the low number of parameters required. As an example, we cite the uncoupled model proposed by Bernard and Remond (2012) who introduced the concept of multiplying the parameter representing the growth rate at the optimal temperature by a temperature-dependent function ϕ_T bounded between 0 and 1, being 1 the value assumed at the optimal temperature. This brief model overview showed the different issues/approaches used in literature to describe the key aspects of microalgae growth (light, temperature, nutrients). The majority of the above-cited works aimed at modelling ideal growth conditions (through lab-scale experimental tests) to focus on the fundamental bioprocesses involving microalgae growth. Nonetheless, modelling microalgae cultivation at large scales introduce further complications, such as the description of light and nutrients distribution along the reactor depth, the representation of heat and mass transfer in the culture medium, the implementation of fluctuating environmental conditions (in particular for open systems), the necessity to describe the hydrodynamics/mixing efficiency with fluid dynamics computation software. For example, Béchet et al. (2011) proposed a universal temperature model applicable to open pond systems and Ali et al. (2017) recently presented a novel empirical correlation to estimate the heat transfer in raceway ponds for different pond sizes and depths. Heat transfer in outdoor raceway ponds was modeled with the effects of pond design, hydrodynamics, and environmental conditions. Park and Li (2015) developed a CFD model to show the variability of biomass at different locations within the system, as well as the light attenuation dependent on depth and cell concentration. Then, it showed that biomass productivity was significantly affected by changes in the incoming CO_2 concentration, while the paddlewheel velocity had no significant effect under turbulent conditions. Another modelling approach was proposed by Solimeno et al. (2015); the aim of the study was to calibrate a mechanistic model built in COMSOL Multiphysics ^{TM} platform to describe both photorespiration, photochemical quenching and photoinhibition together with the influence of temperature, light intensity, pH and nutrients on microlgae growth. In order to assess the mixing effects in a complex hydrodynamic regime Hartmann et al. (2014) proposed a model to reconstruct the light profile received by a single cell. It simulated cells Lagrangian trajectories and the light distribution, hence modelling the light pattern perceived by a cell. This pattern was finally used with a dynamical model for photosynthesis in order to estimate the average growth over a set of trajectories.

In conclusion, the available literature offers a huge quantity of models based on several assumptions on different limiting factors. For this reason, we need to admit that still a lot of key phenomena related both to photosynthesis/metabolism and to cultivation scaling from lab conditions to large-scale production have to be assessed.

1.4 Motivation of the work

Microalgae are one of the most promising renewable feedstocks to cope with fossil sources depletion. Despite their potential, the lack of perfect knowledge about both the fundamental biomass growth processes and the actual possibility to achieve economically satisfactory productivity at large-scale currently raises the question of the real effectiveness of this technology. Mathematical modelling can be of great help to bridge the current gap of uncertainty on microalge growth behavior. Models are in fact valuable tools to support both microscale biological processes representation and macroscale cultivation design and management, with consequent potential increase of cultivation profitability. This Thesis aims at exploring some modelling issues considering both a microscopic and macroscopic perspective. In particular, the following areas have been tackled:

- Microscale modelling. We considered the extension of an existing mathematical state model validated through growth and fluorescence data. In fact, the current model is limited to PSII state description. In this Thesis, we proposed a modelling approach to link PSII dynamics to the whole electron chain transport, in order to describe the dynamics of all the protein complexes involved in the process.

- Macroscale modelling. Open pond systems are commonly considered less efficient than PBRs in terms of productivity; nonetheless open pond systems are commonly operated at constant dilution rates and by maintaining constant pond depth for all the cultivation duration. A new model, developed by coupling existing growth/temperature sub-models with real meteorological data, was used to simulate and optimize an open pond system by acting on independent inflow and outflow rates manipulation, hence allowing depth variation. The utilization of this dynamic model will underline the benefits of model building activities on practical process optimization and control.

In addition to the above, some investigations on MBDoE techniques were conducted. Biological tests for model calibration/validation are often based on fixed protocols (often generated by shared experience or literature reviews) with the risk to conduct sub-informative experimental tasks. In order to cope with this problem, model-based design of experiments (MBDoE) techniques represent a valuable tool for the rapid assessment and development of mathematical models at different levels of the model building procedure, in order to reduce time and costs required for model identification. For this reason, a modified version of classic MBDoE procedure based on online redesign of experiments was proposed, with the perspective to be used for long duration experiments on microalge growth. Since the proposed methodology was not tested directly on microalgae models, it will be discussed in a separate Appendix.

1.5 Dissertation roadmap

A scheme of the Thesis roadmap is presented in Fig.1.6. It presents the main structure of the Thesis, split into two macro-regions associated to the two different approaches used to investigate microalgae growth dynamics.

Chapter 2 presents an extension of the semi-mechanistic model proposed by Bernardi et al. (2016). This state model describes the main biological processes acting on PSII in time scales from milliseconds (photoproduction) to minutes/hours/days (photoinhibition, non-photochemical quenching, photoacclimation). The model was developed and validated by coupling classical photosynthesis rate samples to fast pulse amplitude modulation (PAM) flourescence experiments. The extended model aims at describing the dynamics of the key

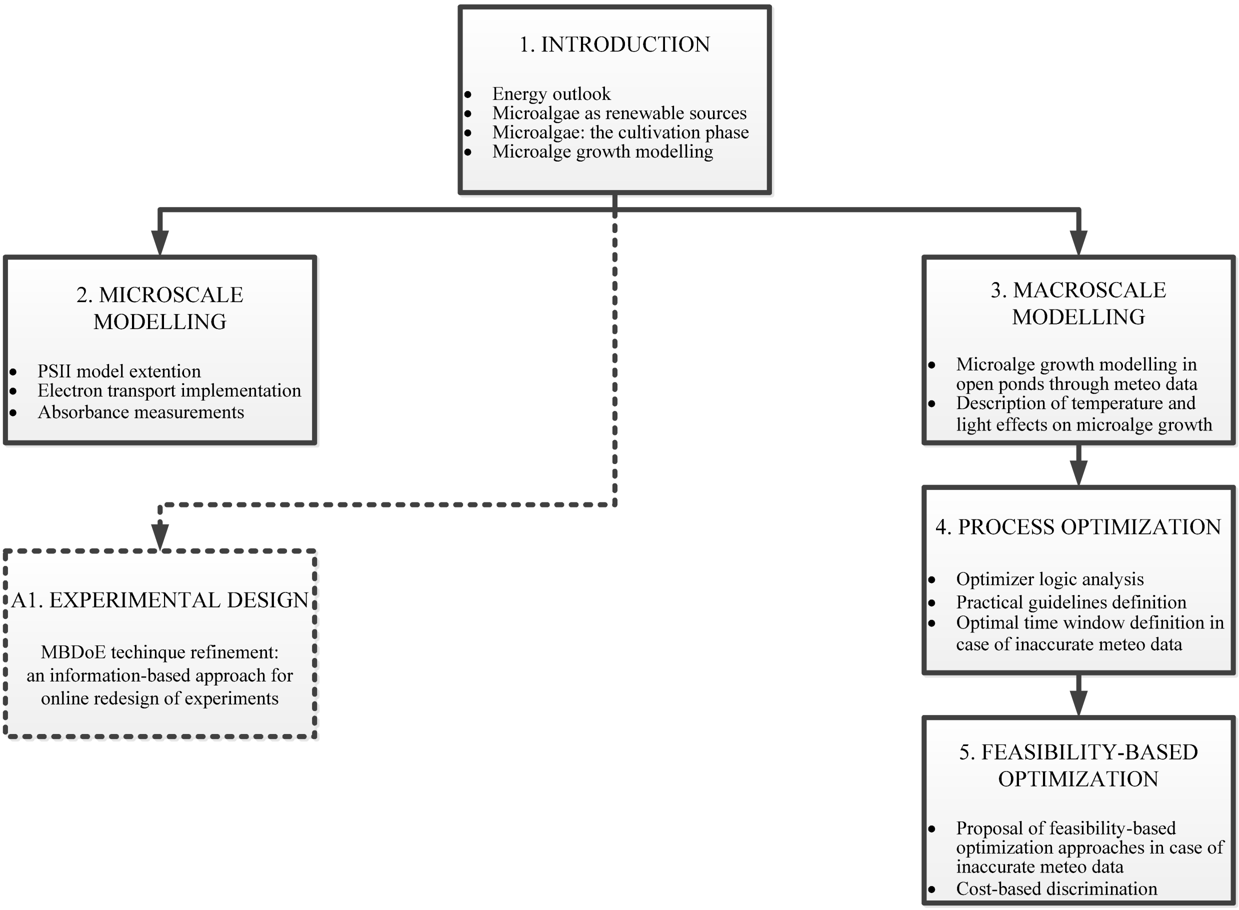


Figure 1.6: *Thesis content roadmap*

protein complexes involved in the electron transport chain through absorbance data.

Chapter 3 presents a detailed description of a new dynamic model describing microalgal productivity in outdoor open ponds; this model was generated by coupling existing sub-models predicting temperature fluctuation effects on the cultivation system (and on growth parameters, too) and the light distribution impact on microalgae growth to real weather data forecasts. The proposed model has been used to perform different optimization tasks in order to check the benefits of acting on inlet and outlet flowrates dynamic manipulation with respect to standard management operation (constant residence time and depth).

Chapter 4 aims at investigating the optimization strategies proposed at different seasons and deriving a unique reduced set of 'rules of thumb' to be used for future practical application. The analysis was conducted both assuming perfect weather forecasts and uncertain meteo data.

Chapter 5 presents a feasibility study to prevent productivity loss and potential critical conditions caused by wrong control manipulation due to inaccurate weather predictions. Two different approaches were proposed, compared and discriminated in terms of productivity, water demand and related process costs.

Chapter 6 summarizes the main achievements reached in this Thesis and proposes some hints for future work.

In Appendix A a new online model-based design of experiments technique to exploit the progressive increase of information resulting from running experiments in case of high parametric mismatch is discussed. The proposed technique is demonstrated with two simulated case studies not strictly related to the Thesis topic. Nonetheless, the results obtained are reported here due to their potential for microalgae growth models calibration/validation.

Chapter 2

A semi-mechanistic model to describe photosynthetic electron transport

In this Chapter a dynamic model predicting electron transport in thylakoid membranes is described in detail. The proposed model is based on the existing model by Bernardi et al. (2016) that couples PSII dynamics representation with fluorescence fluxes quantitative description.

2.1 Introduction

Despite microalgae potential as renewable feedstock, one of the main issues to address is bridging the gap between maximal theoretical productivity and practical biomass productivity in large scale cultivation systems. To meet this objective, a quantitative description of key phenomena affecting microalgae growth, such as light utilization, can be useful to improve our current knowledge for future process optimisation. Several modelling approaches have been proposed over the recent years, as described in the Introduction of this Thesis. The focus here is on the so-called state models, frequently used in literature for their efficiency in describing photosynthetic operation. Nikolau et al. (2015) have recently proposed a semi-mechanistic model (based on Han (2002) 'state-model' approach) to link three distinct processes acting on PSII at different time scales (photoproduction, photoinhibition and photoregulation (NPQ)) to the respective fluorescence fluxes.

Nikolau et al. (2015) showed how considering chlorophyll-a fluorescence dynamics may help providing reliable predictions of the photosynthetic response under variable light conditions, thus allowing for key photosynthetic mechanisms mathematical modelling.

However, this fluorescence model was not able to describe photoregulation adequately over long-term experiments, thus advocating for a more detailed biological representation.

Recently, Bernardi et al. (2016) have considered the model by Nikolau et al. (2015) and concluded that photoregulation mechanisms were based on at least two interdependent processes, hence requiring a more complex representation of NPQ dynamics and, consequently, more experimental data for model calibration. Bernardi et al. (2016)

proposed MBDoE techniques to provide a systematic approach to determine the different experimental protocols that would maximise the avaiable information content for model identification. Eventually, a complete description of the PSII was reached, leading to a significantly improved description of photosynthetic production at different light irradiances.

In this Chapter, Bernardi et al. (2016) work has been used as a starting point to describe the photosynthesis-driven electron transport. In section 2.2 a brief description of the model proposed by Bernardi et al. (2016) is presented; in section 2.3 the proposed model extension for electron transport description is showed. In section 2.4 the calibration/validation were presented and discussed, whereas section 2.5 concludes the Chapter and give some hints for future works.

2.2 The PSII model

The model proposed by Nikolau et al. (2015) and Bernardi et al. (2016) is summarized below. The dynamics of PSII is described according to Han (2002) modelling approach, which is based on the assumption of varying photosystem states. In particular, the reaction centers of PSII can be open (electron free, A), closed (electron charged, B) or inhibited (light damaged, C). Each photosystem can change its state, according to the following set of equations:

$$\frac{dA}{dt} = -I\sigma_{PSII}A + \frac{1}{\tau_{PSII}}B \tag{2.1}$$

$$\frac{dB}{dt} = I\sigma_{PSII}A - \frac{1}{\tau_{PSII}}B + k_r C - k_d \sigma_{PSII} I B \tag{2.2}$$

$$1 = A + B + C, \tag{2.3}$$

where I is the light irradiance (μE m^{-2} s^{-1}), σ_{PSII} is the effective cross-section of PSII (m^2 μE^{-1}), τ_{PSII} is the turnover rate (s), k_r is repair rate constant (s^{-1}) and k_d is the damage rate constant (-). The effective cross-section σ_{PSII} is directly correlated to fluorescence through the equation:

$$\sigma_{PSII} = \frac{\sigma_{tc}\eta_P {\phi_f}^A}{4N_{chl}}, \tag{2.4}$$

where σ_{tc} is the total cross-section (m^2 g$_{chl}^{-1}$), η_P is the rate of photoproduction related to the rate of fluorescence $(-)$, ${\phi_f}^A$ is the quantum yield of fluorescence of an open reaction centre of PSII and N_{chl} is the chlorophyll specific number of photosynthetic units (μmol$_{O_2}$ g$_{chl}^{-1}$). The number 4 in the denominator is a stoichiometric factor reflecting the minimum theoretical value of 4 electrons per oxygen molecule produced during water splitting. The quantum yield of fluorescence ${\phi_f}^A$ is given by the following expression:

$${\phi_f}^A = \frac{1}{1 + \eta_P + \eta_D + \eta_{qE}}, \tag{2.5}$$

Table 2.1: *Parameter estimates along with their 95% confidence interval and t-values. The reference t-value is 1.65. Data extracted by Bernardi et al. (2016). Note that parameter S_f assume different values for each experimental tasks used for parameter estimation (in this case, 3 set of experiments have been used.)*

Parameter	Estimated value	95% conf. int.	t-value 95%	Units
ξ_F	2.68×10^{-1}	3.50×10^{-2}	7.67	s^{-1}
ξ_S	1.32×10^{-3}	6.97×10^{-5}	18.88	s^{-1}
I_{qE}	5.95×10^{2}	2.07×10^{1}	28.76	$\mu\mathrm{Em}^{-2}\mathrm{s}^{-1}$
k_d	9.95×10^{-7}	2.67×10^{-7}	3.73	–
k_r	5.10×10^{-5}	2.67×10^{-5}	1.78	s^{-1}
N_{chl}	4.83×10^{-1}	7.52×10^{-2}	6.43	$\mu\mathrm{mol}_{\mathrm{O}_2}\mathrm{g}_{\mathrm{chl}}^{-1}$
n_{qE}	2.40×10^{0}	1.27×10^{-1}	18.87	–
η_I	1.41×10^{1}	3.98×10^{0}	3.54	–
$\bar{\eta}_{qE}^{F}$	5.96×10^{0}	4.98×10^{-1}	11.95	–
$\bar{\eta}_{qE}^{S}$	1.23×10^{1}	5.75×10^{-1}	21.35	–
$\bar{\eta}_{qE}^{C}$	2.47×10^{1}	1.69×10^{0}	14.58	–
η_P	1.04×10^{1}	2.33×10^{-1}	44.54	–
σ_{tc}	7.33×10^{-1}	7.50×10^{-2}	6.84	$\mathrm{m}^2\mathrm{g}_{\mathrm{chl}}^{-1}$
τ_{PSII}	6.95×10^{-3}	7.50×10^{-4}	9.26	s
S_{f1}	1.81×10^{0}	3.01×10^{-1}	6.82	$\mathrm{Vg}_{\mathrm{chl}}\mathrm{m}^{-2}$
S_{f2}	2.06×10^{0}	3.01×10^{-1}	6.81	$\mathrm{Vg}_{\mathrm{chl}}\mathrm{m}^{-2}$
S_{f3}	1.30×10^{0}	1.90×10^{-1}	6.82	$\mathrm{Vg}_{\mathrm{chl}}\mathrm{m}^{-2}$

where η_D is the rate of basal thermal decay related to the rate of fluorescence $(-)$ and η_{qE} is the rate of energy dependent quenching related to the rate of fluorescence $(-)$. Given that a first-order process is not enough to describe the photo regulation mechanism in an accurate way, Bernardi et al. (2016) proposed the following more complex dynamics:

$$\eta_{qE} = \alpha_F(\bar{\eta}_{qE}^{F} + \alpha_S \bar{\eta}_{qE}^{C}) + \alpha_S \bar{\eta}_{qE}^{S}, \tag{2.6}$$

where $\bar{\eta}_{qE}^{F}$, $\bar{\eta}_{qE}^{S}$ and $\bar{\eta}_{qE}^{C}$ represent different rates of NPQ processes, whereas α_F and α_S are, respectively, the fast and slow activities of NPQ, whose dynamics are described through the following equations:

$$\frac{d\alpha_F}{dt} = \xi_F(\alpha_{SS} - \alpha_F) \tag{2.7}$$

$$\frac{d\alpha_S}{dt} = \xi_S(\alpha_{SS} - \alpha_S), \tag{2.8}$$

where ξ_F and ξ_S are the fast and slow time constant for NPQ and α_{SS} is the reference steady state value reached during the light phase, defined as:

$$\alpha_{SS} = \frac{I^{n_{qE}}}{{I_{qE}}^{n_{qE}} + I^{n_{qE}}}. \tag{2.9}$$

Parameters I_{qE} and n_{qE} in the above sigmoid function represent, respectively, the irradiance level at which half of the NPQ is triggered and the sharpness of the switch-like transition. This model has been calibrated and validated through different kind of experiments on

Nannochloropsis gatidana (see Bernardi et al. (2016)). Parameter estimation was performed by using the entity *Parameter estimation* of gPROMS software (4.1 version), which uses the Maximum Likelihood method. The list of the estimated values of the model parameters is reported in Table 2.1. The practical link to measured fluorescence fluxes F_l is modelled as:

$$F_l = S_f \sigma_{tc} \phi_f, \tag{2.10}$$

where S_f is a parameter depending on the characteristics of the PAM fluorometer used and the chlorophyll content of the sample ($\mathrm{V}_{g_{chl}}\ \mathrm{m}^{-2}$) and ϕ_f is the total fluorescence quantum yield ($-$), described with the following expression:

$$\phi_f = \frac{1}{\frac{A}{{\phi_f}^A} + \frac{B}{{\phi_f}^B} + \frac{C}{{\phi_f}^C}}, \tag{2.11}$$

where ${\phi_f}^B$ and ${\phi_f}^C$ are the quantum yield of fluorescence of closed and inhibited reaction centres, respectively. According to the definition of ${\phi_f}^A$, these two new variables are defined as:

$${\phi_f}^B = \frac{1}{1 + \eta_D + \eta_{qE}} \tag{2.12}$$

$${\phi_f}^C = \frac{1}{1 + \eta_I + \eta_D + \eta_{qE}}, \tag{2.13}$$

where η_I is the rate of inhibition related to the rate of fluorescence ($-$).

2.3 Electron transport description

The complete photosynthetic electron transport chain for *Nannochloropsis gatidana* consists of PSII, PQ, Cytb6f, Cytc6, PSI and Fd (Rochaix (2011)). Light-driven charge separation events occur at the level of the above-modelled PSII and PSI, thus generating the electron transport for ATP and NADPH generation. The electron transport is conventionally split into two main electron fluxes (linear (LEF) and the cyclic (CEF) electron flow, see Eberhard et al. (2008)). Some authors also introduced the so-called pseudolinear cyclic electron flow, in order to describe the complex transient behavior of the electron transport chain just after light activation after long dark periods (Vredenberg and Bulychev (2010)). Despite the proposal of this new electron transport flow, the proposed extension model is uniquely based on LEF and CEF description. The proposed model describes the electron transport by simulating the process represented in Fig. 2.1. As you can see, the final electron acceptor complex (Fd) is responsible of the choice of the final electron flow path; in fact, electrons can be either given to final acceptors (not shown in the graph) or re-introduced in the thylakoid membrane through Cytb6f. The related model is therefore based on the following set of equations:

$$\frac{dx_{PQ,1}}{dt} = \frac{1}{\tau_{PSII}} x_{PQ,2} B n_{PSII} - \alpha_{reg} k_{pq} x_{PQ,1} x_{Cytb6f,2} n_{Cytb6f} \tag{2.14}$$

$$\begin{aligned}\frac{dx_{Cytb6f,1}}{dt} &= \alpha_{reg} k_{pq} x_{PQ,1} x_{Cytb6f,2} n_{PQ} - k_{cyt} x_{Cytb6f,1} x_{Cytc6,1} n_{Cytc6} + \\ &+ \alpha_{reg} k_{cyc} x_{Fd,1} x_{Cytb6f,2} n_{Fd}\end{aligned} \tag{2.15}$$

$$\frac{dx_{Cytc6,1}}{dt} = k_{cyt} x_{Cytb6f,1} x_{Cytc6,2} n_{Cytb6f} - k_{cyt} x_{Cytc6,1} x_{PSI,2} n_{PSI} \tag{2.16}$$

$$\frac{dx_{PSI,1}}{dt} = k_{cyt} x_{Cytc6,1} x_{PSI,2} n_{Cytc6} - \sigma_{PSI} I x_{PSI,1} x_{Fd,2} n_{Fd} \tag{2.17}$$

$$\frac{dx_{Fd,1}}{dt} = \sigma_{PSI} I x_{PSI,1} x_{Fd,2} n_{PSI} - k_{cyc} x_{Fd,1} \tag{2.18}$$

$$x_{i,1} + x_{i,2} = 1, \tag{2.19}$$

where each term $x_{i,j}$ refers to the molar fraction of the i-th complex (PQ, Cytb6f, Cytc6, PSI or Fd) at its j-th redox state (j=1 for reduced states, and j=2 for oxidised states). In practice, no inhibited state is assumed for the protein complexes except for PSII. In a similar way, each term n_i refers to the molar quantity of the i-th complex. In particular, Eq.2.14 describes PQ redox state dynamics: the first right-hand term represents PQ reduction by light excited PSIIs, whereas the second right-hand term represents PQ oxidation due to electron transport to oxidised Cytb6f centers. Then, Eq.2.15 describes Cytb6f dynamics: the first right-hand term represents Cytb6f reduction by reduced PQs, the second right-hand term represents Cytb6f oxidation due to electron transport to oxidised Cytc6 centers and the third right-hand term represents Cytb6f reduction due to cyclic electron flow. Eq.2.3 represents Cytc6 dynamics: the first right-hand term represents Cytc6 reduction by reduced Cytb6fs and the second right-hand term represents Cytc6 oxidation due to electron transport to oxidised PSI centers. PSI dynamics is given by Eq.2.17; in this case the first right-hand term represents PSI reduction by reduced Cytc6 electron carriers and the second right-hand term represents PSI oxidation due to light-driven transport to Fd. Finally, Eq.2.18 describes Fd state dynamics: the first right-hand term represents PSI driven reduction of Fd whereas the second term represents the total electron transport to final acceptors. In other words, no electron transport between two generic complexes A and B is allowed if B is reduced or A is oxidised. The link between this model extension and Bernardi et al. (2016) work is given by the implementation of parameter τ_{PSII} and B in the model equations. k_{pq}, k_{cyt} and k_{cyc} are the different electron transport rates between the electron carriers all along the electron transport chain. The variable α_{reg}, introduced to simulate the regulation mechanism used by the cell to control the electron flux, was modeled through the following equation:

$$\frac{d(1-\alpha_{reg})}{dt} = \xi_{reg}(\beta_{reg,SS} - (1-\alpha_{reg})); \qquad \beta_{reg,SS} = \eta_{reg}\frac{I}{I+1}; \tag{2.20}$$

The proposed formulation is based on the assumption that transport reduction at steady state ($\beta_{reg,SS}$) practically does not depend on light intensity magnitude; the only effect of light is triggering the regulation mechanism. The parameters η_{reg} and ξ_{reg} represents the efficiency and the activation constant of the regulation mechanism, respectively. The α_{reg}

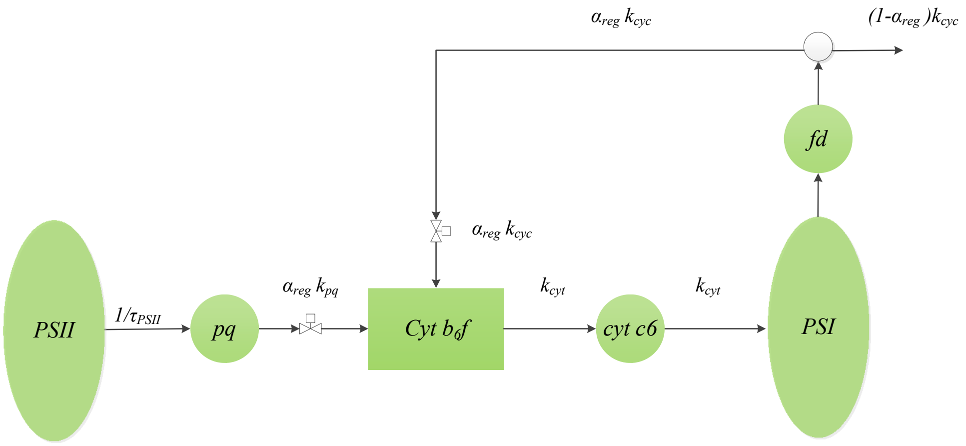


Figure 2.1: *Model flowsheet. The arrows show the connections (and the related kinetics) between the protein complexes involved in the electron transport.*

coefficient was used to simulate the cyclic electron transport decrease after actinic light activation. In this Thesis work two different hypothesis were considered and compared:

- *First case-With Acidification.* The electron transport between PQ and Cytb6f at light conditions is gradually reduced due to lumen acidification caused by cyclic electron transport high rates (see Tikhonov and Vershubskii (2014)). The coefficient α_{reg} was then used to simulate both cyclic electron transport decrease after the initial transient and the linear electron transport reduction due to lumen acidification.

- *Second case-Without Acidification.* The electron transport between PQ and Cytb6f at light conditions is unaffected by lumen acidification (in mathematical terms, the α_{reg} coefficient in Eq.2.14 is removed).

The following assumptions were used for both the two case studies: *(i)* the electron transport between two complexes only takes place where the two complexes have opposite redox state; as an example, reduced PQ complexes can react uniquely with oxidised Cytb6f complexes. This behavior has been highlighted in the previous equations by coloring the subscripts of each protein complex fraction in a different way (red for oxidised complexes and blue for reduced ones); *(ii)* Cytc6 was assumed to be non-zero at the beginning of the light phase. This behavior could be explained by considering that at the end of any light treatment, the excess of electrons is blocked just before PSI (since no oxidation takes place at dark conditions); *(iii)* we assumed the following stoichiometry for the quantity of each complex involved in the electron flow:

- n_{PSI} was set equal to 1 mol;
- $n_{PSI} = 1.5\; n_{PSII}$ (Falkowski and Raven (2007));
- $n_{PQ} = 30\; n_{PSII}$ (Falkowski and Raven (2007));

- $n_{Cytb6f} = 1.5\ n_{PSII}$ (Falkowski and Raven (2007)).

The values of n_{cytc6}/n_{PSII}, n_{fd}/n_{PSII}, $\sigma_{PSII}/\sigma_{PSI}$ have been estimated due to the low information on these parameters in the literature.

2.4 Results and discussion

In the following sub-paragraphs the calibration and validation tasks are defined and discussed; it should be noted that the model parameters of Bernardi et al. (2016) model are kept constant to the original values reported in Table 2.1. The parameters considered in the following are therefore uniquely related to the model extension proposed in this study.

2.4.1 Model calibration

The microalgae strain *Nannochloropsis gaditana* (CCAP, strain 849/5) was grown in a sterile, filtered F/2 medium, using sea salts (32 $\mathrm{gL^{-1}}$) from Sigma, 40 mMTris HCl, pH 8 and Sigma Guillard's (F/2) marine water enrichment solution. Growth experiments were performed in the multi-cultivator MC 1000-OD system (Photon Systems Instruments, Czech Republic) at a temperature of 21°C and a light intensity of 100 $\mu\mathrm{E\,m^{-2}s^{-1}}$ provided continuously by an array of white LEDs. The suspension culture was constantly mixed and aerated by bubbling air. Pre-cultures were grown at 100 $\mu\mathrm{E\,m^{-2}\,s^{-1}}$ in glass bottles of 0.25 L under a continuous airflow, enriched with 5% CO_2. After reaching the exponential growth phase, the pre-culture was centrifuged and re-suspended in fresh medium to have a final concentration of $9 \cdot 10^6 \mathrm{cells\,mL^{-1}}$, before introduction in the multi-cultivator. The culture analyzed was kept in exponential phase by dilution with fresh medium. Spectroscopic analyzes were performed in vivo using a Joliot-type spectrophotometer (JTS-10, Biologic, France). The spectroscopic quantification was performed by measuring the behavior of the primary electron donor (P700) at 705 nm in intact cells at a final concentration of $300 \cdot 10^6$ cells $\mathrm{ml^{-1}}$. The experimental task was conducted by exposing the samples to different actinic lights (from limiting to saturating actinic light; 80-150-320-940-2050 mol of photons $\mathrm{m^{-2}\ s^{-1}}$) for 15000 ms to maximize P700 donor oxidation at a given actinic light and reach a steady state. At the end of each light treatment, the light was switched off for the oxidised P700 re-reduction to occur subsequently in the dark.

The measurement variable used for models calibration is the oxidised fraction of PSI centers ($x_{PSI,2}$), obtained as the ratio between the absorbance signal at normal conditions at a specific light (80-150-320-940-2050 mol of photons $\mathrm{m^{-2}\ s^{-1}}$) and the absorbance signal obtained at 2050 mol of photons $\mathrm{m^{-2}\ s^{-1}}$ through cell-pretreating with DBMIB-poisoner, corresponding to complete oxidation of all PSI centers. Parameter estimation was performed by using the entity *Parameter estimation* of gPROMS software (4.1 version), which uses the Maximum Likelihood method. The lists of the obtained parametric estimates for the two case studies are reported in Tables 2.2 and 2.3. The results showed that for both the

Table 2.2: *First case-With acidification. Parameter estimates along with their 95% confidence interval and t-values. The reference t-value is 1.65.*

Parameter	Estimated value	95% conf. int.	t-value 95%
ξ_{reg}	4.22	0.48	8.77
$x_{cytc6,1}(t=0)$	0.22	0.036	6.04
k_{cyc}	158.66	34.40	4.61
k_{cyt}	17.13	2.67	6.41
k_{pq}	22.56	9.54	2.37
η_{reg}	0.96	0.018	51.48
n_{cytc6}/n_{PSII}	6.06	1.18	5.14
n_{fd}/n_{PSII}	0.27	0.04	5.53
$\sigma_{PSII}/\sigma_{PSI}$	0.56	0.15	5.01

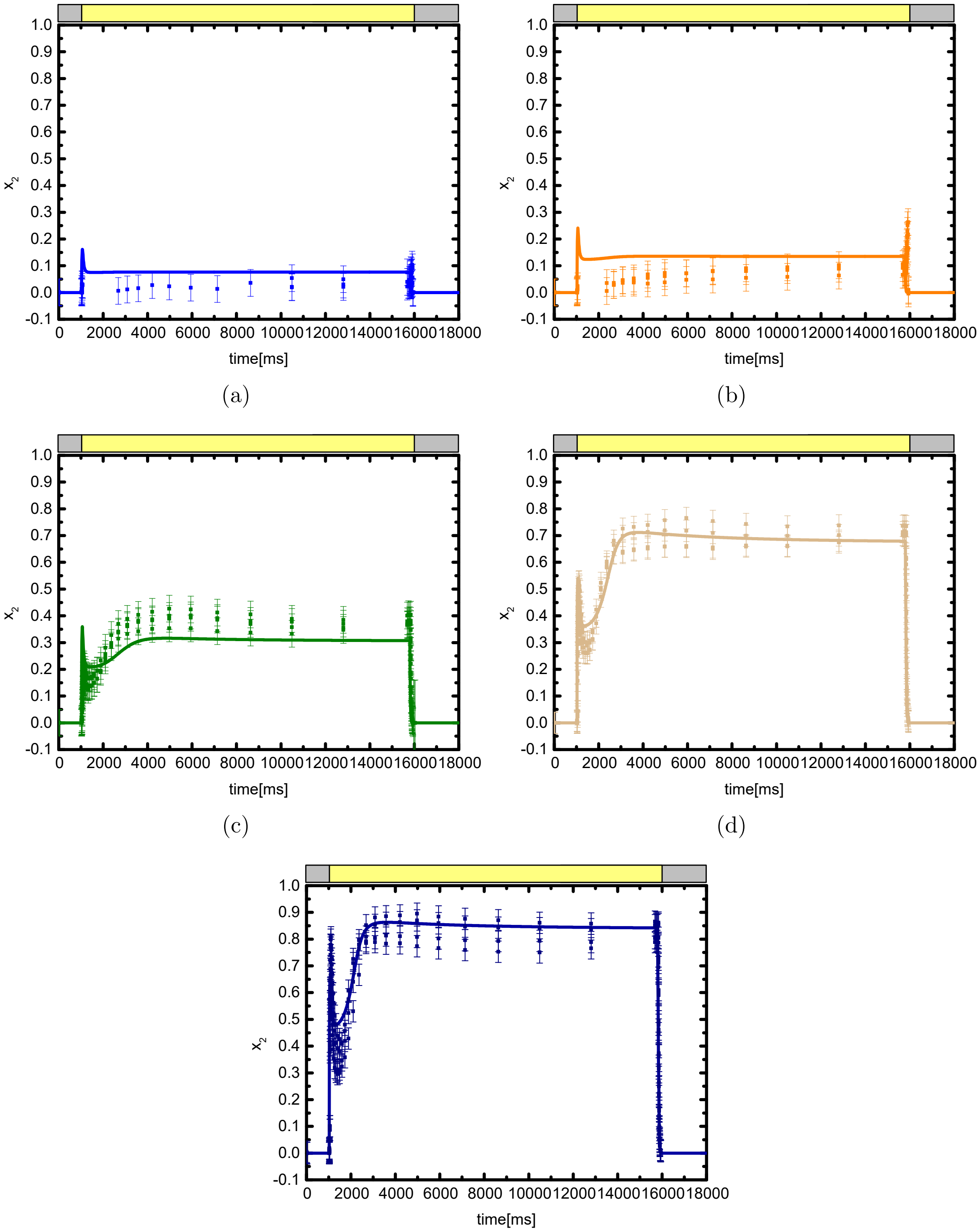
Table 2.3: *Second case-Without acidification. Parameter estimates along with their 95% confidence interval and t-values. The reference t-value is 1.65. The asterisks highlight the parameters whose estimates assume the preset physical boundary values: in this cases no statistical info are provided.*

Parameter	Estimated value	95% conf. int.	t-value 95%
ξ_{reg}	1.33	0.07	18.06
$x_{cytc6,1}(t=0)$	1*	–	–
k_{cyc}	11.80	0.90	12.29
k_{cyt}	130.78	15.14	8.64
k_{pq}	0.95	0.04	24.06
η_{reg}	1*	–	–
n_{cytc6}/n_{PSII}	1.68	0.21	8.03
n_{fd}/n_{PSII}	3.28	0.28	11.81
$\sigma_{PSII}/\sigma_{PSI}$	5.23	0.36	14.44

two case studies all the parameters were estimated in a statistically satisfactory way. All t-values are, in fact, always higher than the reference t-values at 95% of confidence. The case *Without Acidification* presents higher t-values, hence a more precise parameter estimation. Nonetheless, the final estimates of parameters η_{reg} and $x_{cytc6,1}(t = 0)$ (highlighted with asterisks in Table 2.3) tend to overcome the biological threshold and get stuck to the preset boundary values, hence suggesting that *Without Acidification* case study can be based on wrong assumptions. Both the two cases presents a satisfactory goodness of fitting, as reported in Fig.2.2. For the sake of simplicity, only the model description of oxidised x_{PSI} fraction dynamics at each actinic light for the case *With Acidification* is reported in this graph. In fact, the case *Without Acidification* practically showed the same behavior, except for a slightly lower re-oxidation rate at 320 mol of photons m^{-2} s^{-1}.

2.4.2 Model validation

In this paragraph a model validation activity for the above-defined case studies discrimination is described. In this case, absorbance measurements have been used to quantify linear and cyclic electron flows. The experimental values were compared to the simulations obtained



(a) (b)

(c) (d)

(e)

Figure 2.2: *Photosystem 1 oxidized fraction at different actinic lights: (a) 80 $\mu Em^{-2}s^{-1}$, (b) 150 $\mu Em^{-2}s^{-1}$, (c) 320 $\mu Em^{-2}s^{-1}$, (d) 940 $\mu Em^{-2}s^{-1}$, (e) 2050 $\mu Em^{-2}s^{-1}$. The bar above the graphs represents the dark phase (grey) and the light phase (yellow) of the protocol.*

with the parameter values obtained after calibration.

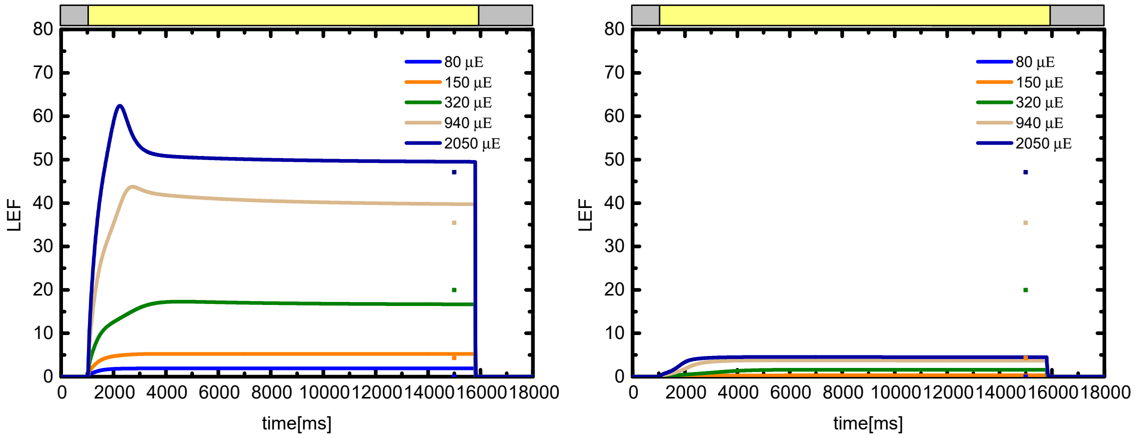
The spectroscopic quantification of the electron flow through the photosynthetic electron transport chain was performed by measuring the behavior of the primary electron donor (P700) to PSI at 705 nm in intact cells at a final concentration of $300 \cdot 10^6$ cells ml^{-1}. The total electron flow TEF obtained at different light treatment was estimated measuring the oxidised P700 re-reduction rates after each illumination in untreated cells. By multiplying this rate constant by the fractions of oxidised P700 at a given light (obtained by comparison with DCMU and DBMIB-poisoned cells treated with 2050 mol of photons m^{-2} s^{-1}, i.e., maximal level of P700 oxidation) we evaluate the number of electrons transferred per unit of time electron flux (Meneghesso et al. (2016)). The same procedure was repeated in samples treated with DCMU (80 M) to evaluate the contribution of cyclic electron flow (CEF) and with DCMU in combination with DBMIB (300 M) to measure any possible residual electron injection into PSI. LEF data was then calculated as the difference between TEF and CEF. The experimental data were compared to the simulated LEF and CEF transport flows, defined as:

$$LEF = (1 - \alpha_{reg}) k_{cyc} x_{Fd,1} x_{PSI,2}, \tag{2.21}$$

$$CEF = (\alpha_{reg}) k_{cyc} x_{Fd,1} x_{PSI,2}. \tag{2.22}$$

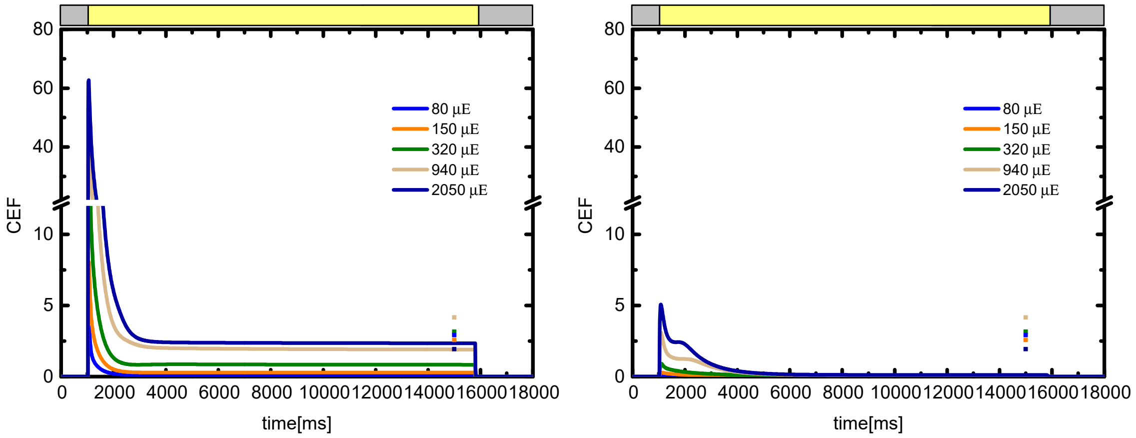
The results, reported in 2.3 and 2.4 show how only the first case (*With Acidification*) is able to represents both LEF and CEF in a good way. The case *With Acidification* shows, in accordance with the experimental data, how LEF behavior dramatically changes with light. In fact, at high actinic lights (940 and 2050 mol of photons m^{-2} s^{-1}) a rapid function increase is shown, then LEF gradually decreases to reach lower stationary values do to acidifcation. At low lights, LEF function is monotone all along the protocol duration. LEF shows an high peak and then a rapid decrease to low values at all lights (the maximum value at steady state is equal to 4.3 e- s^{-1} PSI^{-1}). Conversely, the case *Without Acidification* shows both very low LEFs at all lights and almost null CEF after a low transient peak.

In Figure 2.5 the dynamic simulation of all the protein complexes involved in the electron transport chain is reported. The graphs show that PQs are almost completely reduced after a few milliseconds at each light intensity and this condition is maintained during all the light phase. When light is turned off the PQ centers relax to different redox states depending on the actinic light used. The two cytochromes Cytb6f and Cytc6 start from complete (or almost complete) oxidation to reach different stationary values by about 3s of light, after a transient peak (whose magnitude is strongly dependent on the actinic light used). When light is turned off, both Cytb6f and Cytc6 are completely reduced. Finally, Fd starts from zero reduced centers, shows a light-dependent peak after turning on the actinic light, then rapidly decrease to a stationary value. A final complete re-oxidation of all the centers happens when light is turned off. In summary, the current simulations show that the final state of some protein complexes does not correspond to their initial redox state; nonetheless, since low information is given in the literature about electron carriers



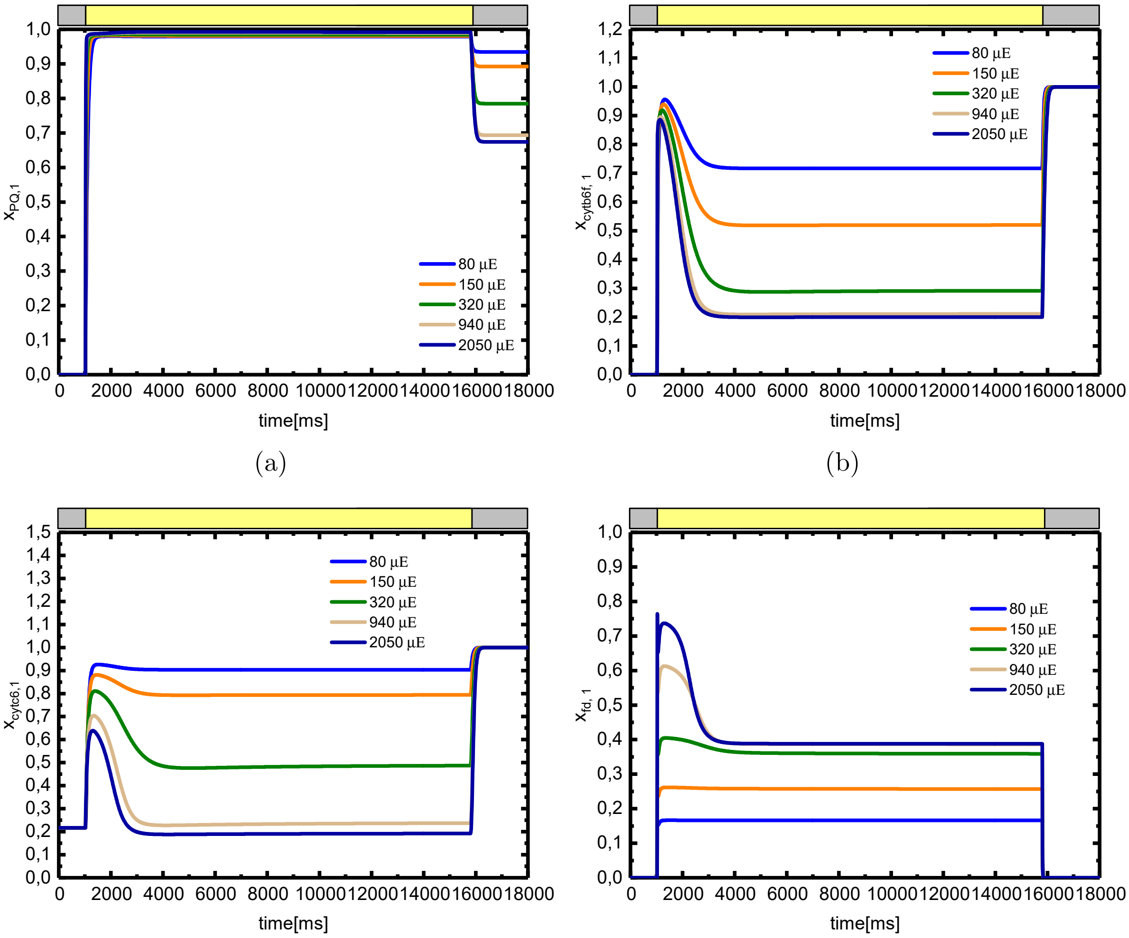
(a) (b)

Figure 2.3: *LEF at different actinic lights (80, 150, 320, 940, 2050 $\mu Em^{-2}s^{-1}$) for case study: (a) with lumen acidification and (b) without lumen acidification. The bar above the graphs represents the dark phase (grey) and the light phase (yellow) of the protocol.*



(a) (b)

Figure 2.4: *CEF at different actinic lights (80, 150, 320, 940, 2050 $\mu Em^{-2}s^{-1}$) for case study: (a) with lumen acidification and (b) without lumen acidification. The bar above the graphs represents the dark phase (grey) and the light phase (yellow) of the protocol.*



(a) (b)

(c) (d)

Figure 2.5: *PQ, Cytb6f, Cytc6, Fd reduced fraction dynamics at different actinic lights (80, 150, 320, 940, 2050 $\mu Em^{-2}s^{-1}$). The bar above the graphs represents the dark phase (grey) and the light phase (yellow) of the protocol.*

dynamics, the assumption of unmodelled dark-phase hysteresis mechanisms for electron carriers relaxation cannot be rejected a-priori. New experiments should therefore be designed to investigate this behavior. Further experimental validation may consists on modifying the current protocol in order to check microalgae behavior at multiple and rapid dark/light cycles and compare it with current model predictions. Furthermore, the existence of dark-phase dependent mechanisms must be investigated and implemented in the existing model to deal with eventually unconsidered mechanisms.

2.5 Conclusions

The proposed extension of Bernardi et al. (2016) work gives a more detailed view of the complete transport chain dynamics. The results show that the model is able to predict both PSI behavior at different light and linear and cyclic flow dynamics. Furthermore, the current model can benefit of fast fluorescence and absorbance measurements, hence reducing the required experimental efforts for model calibration/validation tasks. We precise that the proposed model extension is uniquely meant to describe the photosynthetic process in a more detailed way with respect to the original model. No direct benefits in terms of microalgae growth predictability have been considered as primary objective. Although the model parameters are obtained for *Nannochloropsis gaditana* strain, the model remains applicable to other microalgae species. Nevertheless, new recalibration tasks should be designed to estimate the parameters associated to both Bernardi et al. (2016) model and its extension proposed in this Chapter. Furthermore, the protein complexes involved in the electron transport chain may vary from a strain to another. In this case, the model equations should be adapted to the specific case in order to maintain the biological significance of the model. Further work will focus on investigating the behavior of all the electron carriers between PSII and PSI, in order to improve the process description given by the current model and to confirm (or falsify) some of the model predictions that could not be verified with available experimental data.

Chapter 3

A model to describe algal ponds dynamics accounting for future meteorology

In this Chapter[1] a dynamic model predicting microalgal productivity in outdoor open ponds is described in detail. The selected model consists of three validated sub-models predicting: (*i*) the temperature fluctuations affecting open cultivation systems (Béchet et al. (2013)), (*ii*) the light distribution dynamics in the culture medium (Bernard (2011)) and (*iii*) the algal productivity as a function of both temperature fluctuations and light distribution.

3.1 Introduction

Microalgae is currently being investigated as a promising renewable feedstock for biodiesel production (Mata et al. (2010)) due to its various advantages, such as potential high yields, utilization of non arable land and possible integration with wastewater treatment processes (Foley et al. (2011a)). Nevertheless, current process alternatives for microalgae-based biofuels present critical energy hotspots (Molina Grima et al. (2003); Bennion et al. (2015)) and sub-optimal management, hindering large scale production in an economically satisfactory way. The techno-economic assessments on algal cultivation profitability in open pond systems are currently limited by many uncertainties, especially regarding the biomass productivity that can actually be reached at full-scale. For example, the most cited assessments (Wigmosta et al. (2011); Yang et al. (2011); Moody et al. (2014)) reported productivity values ranging within 1.6-31.6 kg m^{-2} yr^{-1}, depending on the evaluation method used. This large variation of productivity values led to significant differences in assessment conclusions regarding the sustainability of full-scale algal cultivation. A lot of models were recently developed to describe microalgae dynamics and predict yields at full-scale by taking into account different cultivation options and microalgae strains (Bernard et al. (2015); Lee

---

[1] Part of this work is reported in the article by De-Luca R., Q. Béchet, F. Bezzo, O. Bernard (2016). Optimal management of raceways using weather forecasts. *IFAC-PapersOnLine*, 49(7):1062-1067.

et al. (2015)). These models are valuable tools to support both system design (Slegers et al. (2013)) and operation optimization (Muñoz-Tamayo et al. (2013)), with consequent potential increase of cultivation profitability. Nevertheless, building model for open ponds systems presents further complications due to the high variation of key-variables (such as light irradiance, and pond temperature) driven by weather conditions fluctuation. Since the timescale associated with algal growth (i.e. doubling time) has the same order of magnitude than the timescale over which weather forecast can be considered reasonably accurate (i.e. 1-2 days), the knowledge of the weather forecast can be used to optimize system operation on a daily basis and therefore maximize algal productivity. Then, a new model was built to predict both temperature fluctuations in an outdoor pond from weather data and the impact of pond temperature and solar irradiance on algal productivity. The model building task was done by coupling the existing model of microalgae growth proposed by Bernard (2011) with the temperature model of Béchet et al. (2013). The coupled models were then slightly modified by introducing a universal function representing the temperature effects on all the growth-related parameters. The resulting model is described in the following paragraphs.

3.2 The model

A schematic representation of the main structure of the model (coupling biological, thermal and meteorological equations) and the objective function implemented in this work is shown in Fig. 3.1. This representation highlights the interconnection variables between the different sub-models enclosed in the main model; for example, the temperature fluctuation effects are linked to microalgae growth through the ϕ_T function (defined in section 3.2.1), whereas solar irradiance H_s data is implemented in microalgae growth description. A detailed description

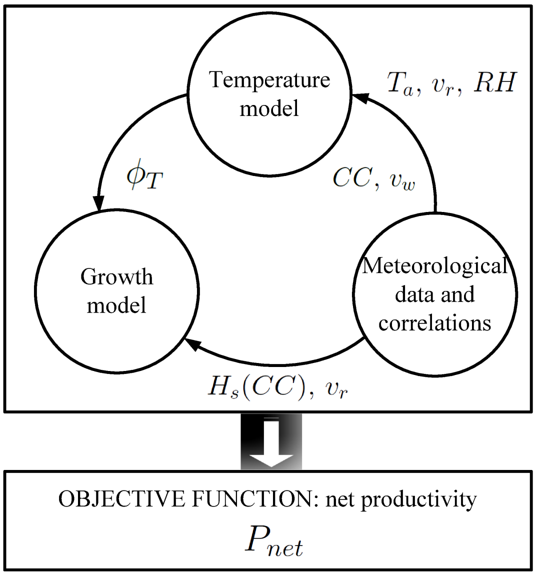


Figure 3.1: *Schematic representation of the models implemented in this work. The arrows show the key variables of interconnection between the various models.*

of all the model equations is presented in sections 3.2.1 and 3.2.2. Section 3.2.3 defines the weather data implemented in the model and the correlations used to simulate the solar irradiance profile. The optimization strategy was applied to algal ponds located in two regions of France representing two different climates. The specific details regarding the optimization at these two locations are discussed in section 3.3. In this section a complete summary of initial/boundary conditions was described, together with the objective function P_{net}, the water demand WD and the 'compensation function' f_{comp}, used to analyze the optimizer behavior discussed in the following Chapter.

3.2.1 Growth model

Let us consider a homogeneous algal open pond of depth l_p (m). If fresh medium is injected into the pond at the rate q^{in} (m^3 s^{-1}) and culture is extracted from the pond at the rate q^{out} (m^3 s^{-1}), the following mass balance can be derived:

$$\frac{d(x_bV)}{dt} = -x_bq^{out} + G(\cdot)V - R(\cdot)V, \tag{3.1}$$

where t is the time variable (s), x_b is the algal biomass concentration (kg m^{-3}), $G(\cdot)$ and $R(\cdot)$ are, respectively, the specific growth and respiration rates (kg m^{-3} s^{-1}), and V is the pond volume (m^3). The open algal raceway pond is supposed to be perfectly mixed (Bernard et al. (2013); Mendoza et al. (2013)). Since the pond is an open system, V varies over time according to the following equation:

$$\frac{dV}{dt} = q^{in} - q^{out} + v_rS - m_eS/\rho_w, \tag{3.2}$$

where S is the pond surface area (m^2), ρ_w is the pond density (kg m^{-3}; assumed equal to water density), v_r is the rainwater flow (m s^{-1}), and m_e is the evaporation mass flux (kg m^{-2} s^{-1}). The specific growth rate $G(\cdot)$ in Eq. 3.1 depends on the biomass concentration x_b, the pond temperature T_p, and the solar irradiance H_s (W m^{-2}). The impact of photoinhibition on microalgae growth was not explicitly included in this study. Indeed, as suggested by Bernard (2011), a Monod kinetics can efficiently represent algal growth at high biomass density. This is explained by the fact that only a small fraction of cells are photo-inhibited in the dense cultures, leading to an average behavior of Monod type. The growth function $G(x_b,\ H_s,\ T_p)$ was therefore expressed as (Béchet et al. (2015a)):

$$G(x_b, H_s, T_p) = \frac{1}{l_p}\int_0^{l_p} \mu_m(T_p)x_b\frac{\sigma_b\eta_H H_s e^{-\sigma_b x_b z}}{K_I(T_p) + \sigma_b\eta_H H_s e^{-\sigma_b x_b z}}dz \tag{3.3}$$

where μ_m is the maximum specific growth rate (s^{-1}), σ_b is the extinction coefficient (set equal to 120 m^2 kg^{-1}), η_H is the fraction of photosynthetically active fraction (PAR) in solar light (set equal to 0.47), z is the local depth (m) and K_I is the half-saturation parameter

(W kg^{-1}). The specific respiration rate $R(\cdot)$ in Eq. 3.1 depends on biomass concentration and pond temperature through the following law (Béchet et al. (2015a)):

$$R(x_b, T_p) = \lambda_r(T_p) x_b, \tag{3.4}$$

where λ_r is the respiration coefficient (s^{-1}). Experimental values for μ_m, K_I and λ_r for *Chlorella vulgaris* were extracted from the study of Béchet et al. (2015a). Bernard and Remond (2012) showed that the evolution of the parameter μ_m with temperature could be fitted to the following function:

$$\mu_m = \mu_{m,max} \phi_T, \tag{3.5}$$

where $\mu_{m,max}$ is the maximum value of μ_m (s^{-1}) and ϕ_T is the temperature-dependent function defined in Bernard and Remond (2012) as follows:

$$\phi_T = \begin{cases} 0 & \text{if } T_p \leq T_{min} \\ \frac{(T_p - T_{max})(T_p - T_{min})^2}{(T_{opt} - T_{min})[(T_{opt} - T_{min})(T_p - T_{min}) - (T_{opt} - T_{max})(T_{opt} + T_{min} - 2T_p)]} & \text{if } T_{min} < T_p \leq T_{max} \\ 0 & \text{if } T_p > T_{max}. \end{cases} \tag{3.6}$$

The function ϕ_T includes three parameters: T_{min}, T_{opt} and T_{max} (°C). T_{min} is the temperature below which the growth is assumed to be zero, T_{max} is the temperature above which there is no growth, T_{opt} is the temperature at which $\mu_m = \mu_{m,max}$. As λ_r and K_I exhibited similar evolution with temperature, the same function ϕ_T was used for fitting the evolution of these two parameters with temperature:

$$K_I = K_{I,max} \phi_T, \tag{3.7}$$

$$\lambda_r = \lambda_{r,max} \phi_T. \tag{3.8}$$

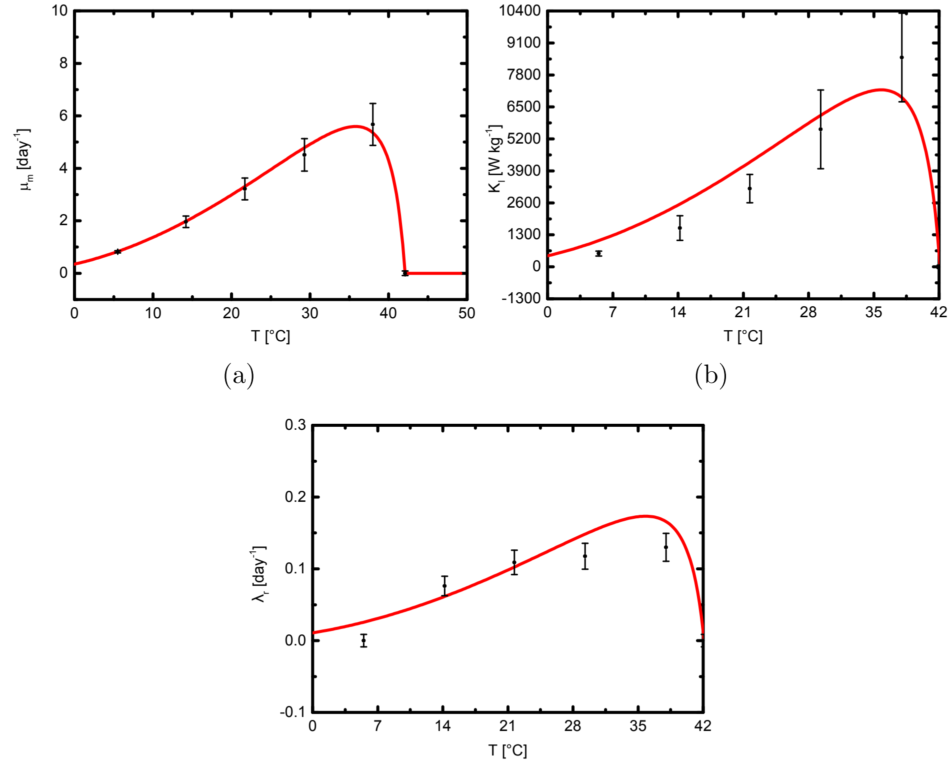
Fitting these parameters was performed by using the entity *Parameter estimation* of gPROMS software (4.1 version), which uses the Maximum Likelihood method. The function ϕ_T successfully fitted the experimental data for both μ_m, K_I and λ_r (see Fig. 3.2). The complete set of fitted values is reported in Table 3.2.

Table 3.1: *Experimental evaluation of growth and respiration parameters at different temperatures (Béchet et al. (2015a)). The reported data of μ_m and λ_r are converted from $kg_{O_2}/kg_{biomass}/s$ to $kg_{biomass}/kg_{biomass}/s$ through the corrective factor (0.76) proposed in the same article.*

Variables	Values					
T_p(°C)	5.5	14.2	21.7	29.3	38.0	42.1
λ_r(s^{-1})	0	$8.82 \cdot 10^{-7}$	$1.26 \cdot 10^{-6}$	$1.36 \cdot 10^{-6}$	$1.51 \cdot 10^{-6}$	0
K_I(W kg^{-1})	540	1580	3180	5590	8510	0
μ_m(s^{-1})	$0.95 \cdot 10^{-5}$	$2.27 \cdot 10^{-5}$	$3.72 \cdot 10^{-5}$	$5.22 \cdot 10^{-5}$	$6.57 \cdot 10^{-5}$	0

Table 3.2: *Temperature-dependent parametric estimates*

Parameter	Physical Definition	Value
T_{min}	Minimum growth temperature	-10.0 (°C)
T_{max}	Maximum growth temperature	42.1 (°C)
T_{opt}	Optimum growth temperature	35.8 (°C)
$\lambda_{r,max}$	Max. respiration coefficient	$2.01{\cdot}10^{-6}$ (s^{-1}) = 0.17 (day^{-1})
$\mu_{m,max}$	Max. specific growth rate	$6.48{\cdot}10^{-5}$ (s^{-1}) = 5.59 (day^{-1})
$K_{I,max}$	Max. half-saturation constant	7192.92 (W kg^{-1})



(a) (b)

(c)

Figure 3.2: *ϕ_T fitting of the experimental data (Béchet et al. (2015a)) related to: (a) μ_m, (b) K_I and (c) λ_r.*

3.2.2 Temperature Model

The temperature dynamics during the cultivation was modeled through a heat balance on the pond that can be expressed as (Béchet et al. (2011)):

$$\rho_w V c_{p_w} \frac{dT_p}{dt} = Q_{ra,p} + Q_{ra,s} + Q_{ra,a} + Q_{ev} + Q_{conv} + Q_{cond} + Q_i + Q_r, \tag{3.9}$$

where c_{p_w} is the specific heat capacity of water (J $kg^{-1}K^{-1}$), $Q_{ra,p}$ is the radiation flow from the pond surface (W), $Q_{ra,s}$ is the total (direct+diffuse) solar irradiance (W), $Q_{ra,a}$ is the radiation flow from the air to the pond system (W), Q_{ev} is the evaporation flow (W), Q_{conv} is the convective flow at the pond surface (W), Q_{cond} is the conductive flow with the ground at the pond bottom (W), Q_i is the heat flow due to the water inflow (W), and Q_r is the heat flow associated with rain (W). This model was validated by (Béchet et al. (2011))

against 1 year of experimental data, collected from a wastewater treatment high rate algal pond; then, the proposed model was compared to other nine models that differed in the formulation of each heat fluxes involved in the temperature dynamics, showing the best fit between the experimental and predicted temperature profiles. The formal expression of each heat flow in Eq. 3.9 was extracted from Béchet et al. (2011). Assuming the water surface to be gray-diffuse, the radiation $Q_{ra,p}$ from the pond surface to the atmosphere was given by Stefan-Boltzmann's fourth power law:

$$Q_{ra,p} = -\varepsilon_w \sigma_{SB} {T_p}^4 S, \tag{3.10}$$

where ε_w is the emissivity of water and σ_{SB} is the Stefan-Boltzmann constant (W m^{-2} K^{-4}). The heat flow associated with solar irradiance $Q_{ra,s}$ was expressed as follows:

$$Q_{ra,s} = (1 - f_a) H_s S, \tag{3.11}$$

where f_a is the theoretical photosynthetic efficiency (the fraction of PAR that is effectively used during the photosynthetic process). The heat flow generated by air radiation was described as:

$$Q_{ra,a} = \varepsilon_a \varepsilon_w \sigma_{SB} {T_a}^4 S, \tag{3.12}$$

where ε_a is the emissivity of the air and T_a is the air temperature (K). The evaporation heat flow Q_{ev} was given by the equation:

$$Q_{ev} = -m_e L_w S, \tag{3.13}$$

where L_w is the water latent heat (J kg^{-1}), whereas m_e is the evaporation rate (kg m^{-2} s^{-1}). The following expression was used to calculate m_e:

$$m_e = K_m \left(\frac{P_w}{T_p} - \frac{RH \cdot P_a}{T_a} \right) \frac{M_w}{R_g}, \tag{3.14}$$

where K_m is the mass transfer coefficient (m s^{-1}), RH is the relative air humidity above the pond surface, M_w is the molecular weight of water (kg mol^{-1}) and R_g is the universal ideal gas constant (Pa m^3 mol^{-1} K^{-1}). P_w and P_a are, respectively, the saturated vapor pressure (Pa) at T_p and T_a, evaluated through the following empirical correlation:

$$P_i = 3385.5 \cdot e^{\left(-8.0929 + 0.97608 (T_i + 42.607 - 273.15)^{0.5}\right)}, \tag{3.15}$$

where the i index represents air or water. The mass transfer coefficient K_m in (3.14) was calculated through the following two correlations:

$$Sh = 0.035 Re^{0.8} Sc^{1/3} \quad \text{for turbulent flows,} \tag{3.16}$$

$$Sh = 0.628 Re^{0.5} Sc^{1/3} \quad \text{for laminar flows,} \tag{3.17}$$

where $Sh = K_m L_{ch}/D_{w,a}$, $Re = L_{ch} v_w/\nu_a$ and $Sc = \nu_a/D_{w,a}$. The dimensionless variables Sh, Re and Sc are, respectively, the Sherwood, Reynolds and Schmidt numbers, L_{ch} is the characteristic pond length (m), $D_{w,a}$ is the mass diffusion coefficient of water vapor in air (m^2 s^{-1}), v_w is the wind velocity (m s^{-1}) and ν_a is the air kinematic viscosity (m^2 s^{-1}). The convective flow Q_{conv}, defined as:

$$Q_{conv} = h_{conv}(T_a - T_p)S \tag{3.18}$$

was calculated by evaluating the heat transfer coefficient h_{conv} (W m^{-2} K^{-1}) value through the following set of correlations:

$$Nu = 0.035 Re^{0.8} Pr^{1/3} \quad \text{for turbulent flows,} \tag{3.19}$$

$$Nu = 0.628 Re^{0.5} Pr^{1/3} \quad \text{for laminar flows,} \tag{3.20}$$

where $Nu = h_{conv} L_{ch}/\lambda_a$ and $Pr = \nu_a/\alpha_a$, λ_a is the air thermal conductivity (W m^{-1} K^{-1}) and α_a the air thermal diffusivity (m^2 s^{-1}). The dimensionless variables Nu and Pr are, respectively, the Nusselt and Prandtl numbers. Meteorological stations measure the wind velocity at a certain height $z_{m,0}$ (m) which usually differs from the height at which v_w is needed in the previous correlations (0.5 m). The conversion of the wind velocity from the height $z_{m,0}$ to 0.5 m was performed by using the following expression:

$$v_w = v_0 \left(\frac{z_m}{z_{m,0}} \right)^{\alpha_{vs}}, \tag{3.21}$$

where v_0 is the wind velocity (m s^{-1}) measured at height $z_{m,0}$ and α_{vs} is a power law exponent. The equation that describes the conductive heat flow between the pond and the soil was based on Fourier's law:

$$Q_{cond} = k_s S \frac{dT_s}{dz}(z = 0), \tag{3.22}$$

where k_s is the soil conductivity (W m^{-1} K^{-1}) and T_s is the soil temperature (K). The value of T_s was obtained from the following equation and initial/boundary conditions:

$$c_{p_s} \rho_s \frac{dT_s}{dt}(z,t) = k_s \frac{d^2 T_s}{dz^2}(z,t) \tag{3.23}$$

$$\begin{cases} T_s(t, z = 0) = T_p(t) & \text{b.c.(1)} \\ T_s(t, z = l_{s_{ref}}) = T_{s_{ref}} & \text{b.c.(2)} \\ \frac{d^2 T_s}{dz^2}(t = 0) = 0 & \text{i.c.} \end{cases} \tag{3.24}$$

where c_{p_s} is the soil specific heat capacity (J kg^{-1} K^{-1}), ρ_s is the soil density (kg m^{-3}), and $T_{s_{ref}}$ is the soil temperature (K) at the reference depth $l_{s_{ref}}$ (m). The heat flow associated with fresh medium inflow Q_i was computed from the equation:

$$Q_i = \rho_w c_{p_w} q^{in}(T_{in} - T_p) \tag{3.25}$$

where T_{in} is the water inflow temperature (K). Finally, the rain heat flow Q_r was expressed as:

$$Q_r = \rho_w c_{p_w} v_r (T_a - T_p) S. \tag{3.26}$$

All the parameters values used in the energy balance are tabulated in the study of Béchet et al. (2011) and reported in Table 3.3.

3.2.3 Meteorological data and correlations

Continuous weather data was linearly interpolated from the 6-hourly weather data extracted from the European Centre for Medium-Range Weather Forecast (ECMWF) website. This weather data was used to determine the dynamics of the air temperature T_a, the sky cloudiness CC, the relative humidity RH, the wind velocity v_w and the rain volumetric flux v_r as all these variables have a significant impact on pond temperature as shown by Béchet et al. (2011). Solar irradiance H_s was computed at any time from both the amount of solar radiation reaching the external surface of the atmosphere at the location considered H_0, and the cloudiness CC. First, H_s was calculated from cloudiness data CC, by using the Kasten and Czeplak correlation (Marthews et al. (2012)):

$$H_s = \begin{cases} 0 & \text{if } \omega < -\omega_s \text{ or } \omega > \omega_s \\ H_c \frac{(4-3(CC/8)^{3.4})}{4} & \text{if } -\omega_s \leq \omega \leq \omega_s, \end{cases} \tag{3.27}$$

where CC is the cloudiness value expressed in okta (range [0−8]), H_c is the clear-sky total irradiance (W m^{-2}), ω is the hour angle which varies from -π to π over 24 hours, -ω_s and ω_s are, respectively, the hour angle values at sunrise and sunset calculated from the expression proposed by Duffie and Beckman (1958): The clear-sky radiation H_c was given by Duffie and Beckman (1958):

$$H_c = H_{d,c} + H_{D,c} = (\tau_{d,c} + \tau_{D,c})H_0 \tag{3.28}$$

where $\tau_{d,c} = H_{d,c}/H_0$ and $\tau_{D,c} = H_{D,c}/H_0$; $H_{d,c}$ and $H_{D,c}$ are, respectively, the diffuse and the direct components of the clear-sky total irradiance (W m^{-2}). The variable H_0 was given by Duffie and Beckman (1958):

$$H_0 = I_{sc}\left(1 + 0.033 \cos \frac{360 N_{day}}{365}\right) \cos\theta_z, \tag{3.29}$$

Table 3.3: *Parameter values of the temperature model*

Parameter	Physical Definition	Value
	Water Parameters	
ρ_w	water density	998 ($\text{kg}\cdot\text{m}^{-3}$)
c_{p_w}	water heat capacity	$4.18\cdot10^3$ ($\text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$)
L_w	water latent heat	$2.45\cdot10^6$ ($\text{J}\cdot\text{kg}^{-1}$)
ε_w	water emissivity	0.97 (-)
M_w	water molecular weight	0.018 ($\text{kg}\cdot\text{mol}^{-1}$)
	Soil Parameters	
k_s	soil thermal conductivity	1.7 ($\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$)
c_{p_s}	soil heat capacity	$1.25\cdot10^3$ ($\text{J}\cdot\text{kg}^{-1}\cdot\text{K}^{-1}$)
ρ_s	soil density	$1.9\cdot10^3$ ($\text{kg}\cdot\text{m}^{-3}$)
$T_{s_{ref}}$	soil temperature at $l_{s_{ref}} = 4.5\text{m}$	286.75 (K)
	Air Parameters	
ε_a	air emissivity	0.8 (-)
ν_a	air kinematics viscosity	$1.5\cdot10^{-5}$ ($\text{m}\cdot\text{s}^{-1}$)
λ_a	air thermal conductivity	$2.6\cdot10^{-2}$ ($\text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$)
α_a	air thermal diffusivity	$2.2\cdot10^{-5}$($\text{m}\cdot\text{s}^{-1}$)
$D_{w,a}$	mass diffusion coefficient of water vapor in air	$2.4\cdot10^{-5}$ ($\text{m}\cdot\text{s}^{-2}$)
	Pond parameters	
V_0	initial pond volume	30 (m^3)
S	pond surface	100 (m^2)
L_{ch}	pond characteristic length	10 (m)
α_{vs}	power law exponent	0.29 (-)
z_m	wind velocity height	0.5 (m)
$z_{m,0}$	wind sensor height	10 (m)
f_a	algal absorption fraction	2.5 (%)
	Universal constants	
σ_{SB}	Stephan-Boltzmann constant	$5.67\cdot10^{-8}$ ($\text{W}\cdot\text{m}^{-2}\cdot\text{K}^{-4}$)
R_g	ideal gas constant	8.314 ($\text{Pa}\cdot\text{m}^3\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$)

where I_{sc} is the solar constant (1367 W m^{-2}), N_{day} is the Julian day of the year and θ_z is the zenith angle (the angle between a vertical axis at the location considered and the sun direction).
The variable $\tau_{D,c}$ was obtained through the following expression (Marthews et al. (2012)):

$$\tau_{D,c} = \exp(-\gamma_{LK} T_{LK} A_{LK}), \tag{3.30}$$

where T_{LK} is an empirical coefficient (set to 2.74). The term γ_{LK} was given by the equation:

$$\gamma_{LK} = \frac{1}{9.4 + 0.9 A_{LK}} \tag{3.31}$$

and A_{LK} is the pressure-corrected air-mass, obtained from the following correlation (Marthews et al. (2012)):

$$A_{LK} = \frac{1}{\cos\theta_z + 0.50572\left(96.07995 - \theta_z \frac{180}{\pi}\right)^{-1.6364}}. \tag{3.32}$$

Finally, the variable $\tau_{d,c}$ is determined through the Erbs correlation (recommended by Duffie and Beckman (1958)):

$$K_d = \begin{cases} 1.0 - 0.09 k_T & \text{if } k_T \leq 0.22 \\ 0.9511 - 0.1604 k_T + 4.388 {k_T}^2 + & \\ -16.638 {k_T}^3 + 12.336 {k_T}^4 & \text{if } 0.22 < k_T \leq 0.8 \\ 0.16527 & \text{if } k_T > 0.8 \end{cases}, \tag{3.33}$$

where $K_d = \tau_{d,c}/(\tau_{d,c}+\tau_{D,c})$ and $k_T = \tau_{d,c}+\tau_{D,c}$.

3.3 The optimization

A detailed description of the initial/physical conditions of the system is given in 3.3.1. Section 3.3.2 defines the objective function for all the optimization tasks performed in the Thesis work. Finally, 3.3.3 and 3.3.4 focus on the definitions of two key variables used for case studies comparisons in the following discussion.

3.3.1 System description

The optimization strategy was applied at two different locations in France (Nice and Rennes) during the first 7 days of different months in 2012. A complete graphical representation of the meteorological data implemented in this Thesis work is reported in Appendix B.
The pond surface S was 100 m^2 and the initial conditions were:

- the initial pond temperature $T_p(t = 0)$ was set at the average value of air temperature $T_{a,avg}$ over the period τ of simulation/optimization;

- the initial biomass concentration x_b was set to 0.4 kg/m^3;
- the initial pond depth l_p was set to 0.3 m;
- the inflow temperature T_{in} was set equal to $T_{a,avg}$.

3.3.2 System optimization

The optimal control strategy aimed at maximizing microalgal productivity, by continuously adjusting pond depth through the injection of fresh medium (q^{in}) or extraction of culture (q^{out}). The control vector $\mathbf{q} = (q^{in},\, q^{out})^{\mathsf{T}}$ was therefore the solution of the following optimal control problem:

$$\begin{array}{ll} \max_{\mathbf{q}} P_{net} = & \max_{\mathbf{q}} \int_0^{\tau} (G(x_b, H_s, T_p) - R(x_b, T_p))V dt \\ \text{s.t.} & \\ & 0 \leq q_i(t) \leq q_{\max} \\ & l_{p,\min} \leq l_p(t) \leq l_{p,\max} \\ & \dot{\boldsymbol{\xi}} = \mathbf{g}(\boldsymbol{\xi}, \mathbf{q}, t), \quad \boldsymbol{\xi}(0) = \boldsymbol{\xi}_0 \end{array} , \tag{3.34}$$

where $\boldsymbol{\xi}$ is the state vector and $q_{\max}$ is the upper bound of the flow rates. The dynamical system $\mathbf{g}(\boldsymbol{\xi},\, \mathbf{q},\, t)$ therefore gathered the biological and thermal dynamics (see Equations (3.1) and (3.9)). The objective function P_{net} is the cumulative difference between growth and respiration during the cultivation period. The chosen formulation is based on the expression proposed in Béchet et al. (2015b) for the specific case where Monod formula is used to represent microalge growth and respiration rate is considered. The main difference between The key difference with the control problem addressed in (Muñoz-Tamayo et al. (2013)) is that the pond depth was made vary in this study. The pond depth was constrained between $l_{p,\min}$ (0.05 m) and $l_{p,\max}$ (0.5 m).

The optimization task was implemented through gPROMS software (4.1 version) by using the default optimization solver NLPSQP, which uses a sequential quadratic programming (SQP) method for the solution of nonlinear programming (NLP) problems. The NLPSQP solver uses **4 tolerances** to decide whether or not a local optimal solution has been reached; the default value of the tolerances was used for all the optimization tasks conducted in this Thesis. For the numerical implementation adopted in this Thesis work, the two control inputs q^{in} and q^{out} were considered as piecewise constant variables within the range [0-1] m^3/s.

In order to assess the gain of productivity obtained with this control strategy, other simulations were performed under 'standard' conditions. These standard conditions correspond to typical assumptions made in the field of micro-algae outdoor cultivation: constant depth of 0.25 m and dilution rate equal to 0.1 day^{-1} (Jorquera et al. (2010), Rogers et al. (2014)).

3.3.3 Water Demand

The net water demand (Water Demand (WD)) associated with algal cultivation was calculated as follows:

$$WD = \int_0^{\tau} q^{in}(t)dt + \max(0, V_0 - V_\tau), \tag{3.35}$$

where V_0 and V_τ are, respectively, the pond volume at the beginning and at the end of the cultivation period. This expression accounts for both water use during fresh water injection into the pond and any net decrease of the pond volume between the start and the end of the cultivation period considered.

3.3.4 The compensation function

The compensation function f_{comp} has often been proposed as a key criterion to be optimized (Takache et al. (2010), Muñoz-Tamayo et al. (2013)). The compensation function is the ratio between the growth and the respiration rates at the pond bottom and it should be as close as possible to 1:

$$f_{comp}(t) = \mu_{m,max} \frac{\frac{\sigma_b \eta_H H_s(t) e^{-\sigma_b x_b(t) l_p(t)}}{K_{I,max}\phi_T(t) + \sigma_b \eta_H H_s(t) e^{-\sigma_b x_b(t) l_p(t)}}}{\lambda_{r,max}}. \tag{3.36}$$

In other words, if the function is higher than 1 the pond productivity could be improved by, for example, increasing the amount of biomass in the system. Conversely, values lower than 1 indicate that the net rate of growth at the pond bottom is negative. Consequently, diluting the system would increase productivity. The optimal biomass concentration $x_{b,opt}$ is therefore reached when f_{comp} equals 1.

3.4 Conclusions

In this Chapter a weather data-based dynamic model predicting microalgal productivity in open ponds has been described in detail. The model describes both temperature fluctuations and light distribution dynamics, hence their effects on microalgae growth efficiency. This model was used to run several optimization tasks in order to check the potential efficiency of the proposed optimization strategy with respect to standard open pond management. We precise that no model mismatch was assumed and that all the results reported in the following are generated by 'in silico' simulations; future experiments should be performed for model validation to estimate the amount of uncertainty generated by each specific sub-model. In the next Chapter, the optimization results obtained with the proposed model will be discussed in detail.

Chapter 4

Exploiting meteorological forecasts for optimal open pond management

The first objective of this Chapter[1] is to investigate how the optimization strategy defined in the previous Chapter can boost algal productivity in outdoor open ponds. The secondary objective is to derive a reduced set of ‘rules of thumb’ which represent ‘optimal operation’ to reduce the time required for future applications of this optimization strategy. We considered both the case where weather forecasts are perfect and the more realistic case where meteorology becomes uncertain after 24 hours.

4.1 Strategy impact on productivity

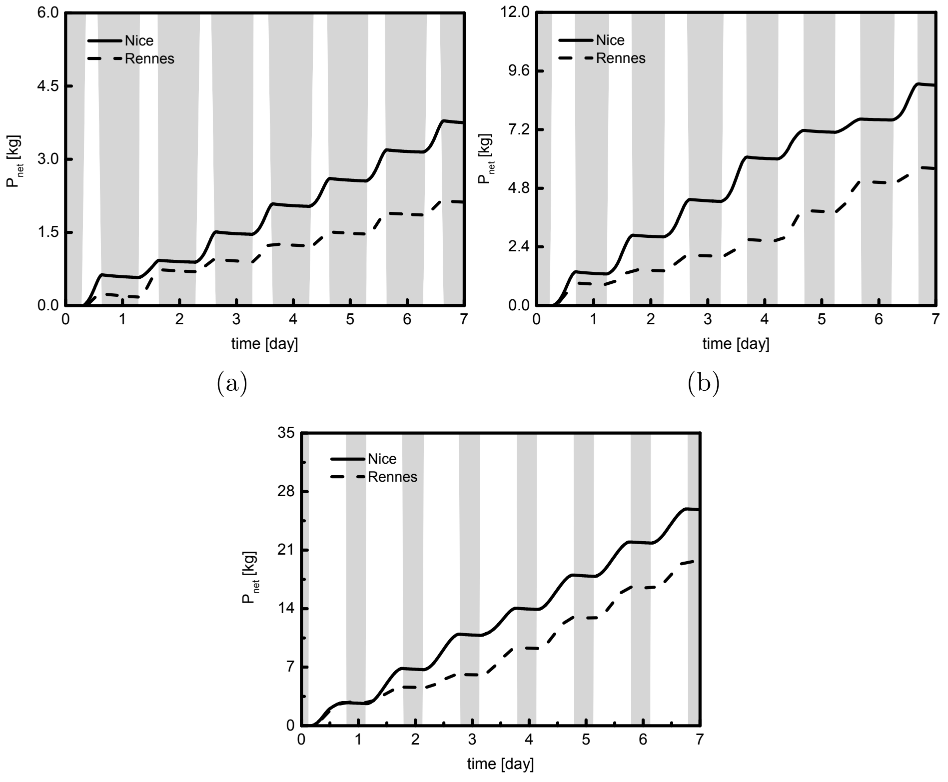
The proposed strategy consists of determining iteratively the optimal inflow and outflow hourly rates for an entire week based on the weekly weather forecasts. Unlike in Muñoz-Tamayo et al. (2013), the culture depth can vary, and thus thermal inertia of the cultivation system can be modified to optimize system temperature fluctuations. Alternatively, we propose an optimization strategy that is based on the knowledge of future weather conditions instead of using a control approach based on on-line measurements. In practice, this optimization task consists of determining the optimal rates at which culture is replaced in the system in advance for an entire week based on the weather forecast for the following week. Table 4.1 shows the algal productivity values obtained during optimized and standard cultivation, at Nice and Rennes and at three seasons (Winter (January), Spring (March), and Summer (July). The reported results show that the optimization strategy investigated in this study significantly increased productivity compared to standard operation, by up to a factor 2.2 for the Summer case in Nice.

---

[1] Part of this work is reported in the article by De-Luca R., F. Bezzo, Q. Béchet, O. Bernard (2017). Exploiting meteorological forecasts for the optimal operation of algal ponds. In: *Journal of Process Control.* (doi: 10.1016/j.jprocont.2017.03.010) and in the submitted article by De-Luca R., F. Bezzo, Q. Béchet, O. Bernard (2017). Meteorological data-based optimal control strategy for microalgae cultivation in open pond systems. In: *Bioresource Technology.*

Table 4.1: *Productivity and water demand: optimal case vs standard pond management (Jorquera et al. (2010), Rogers et al. (2014)).*

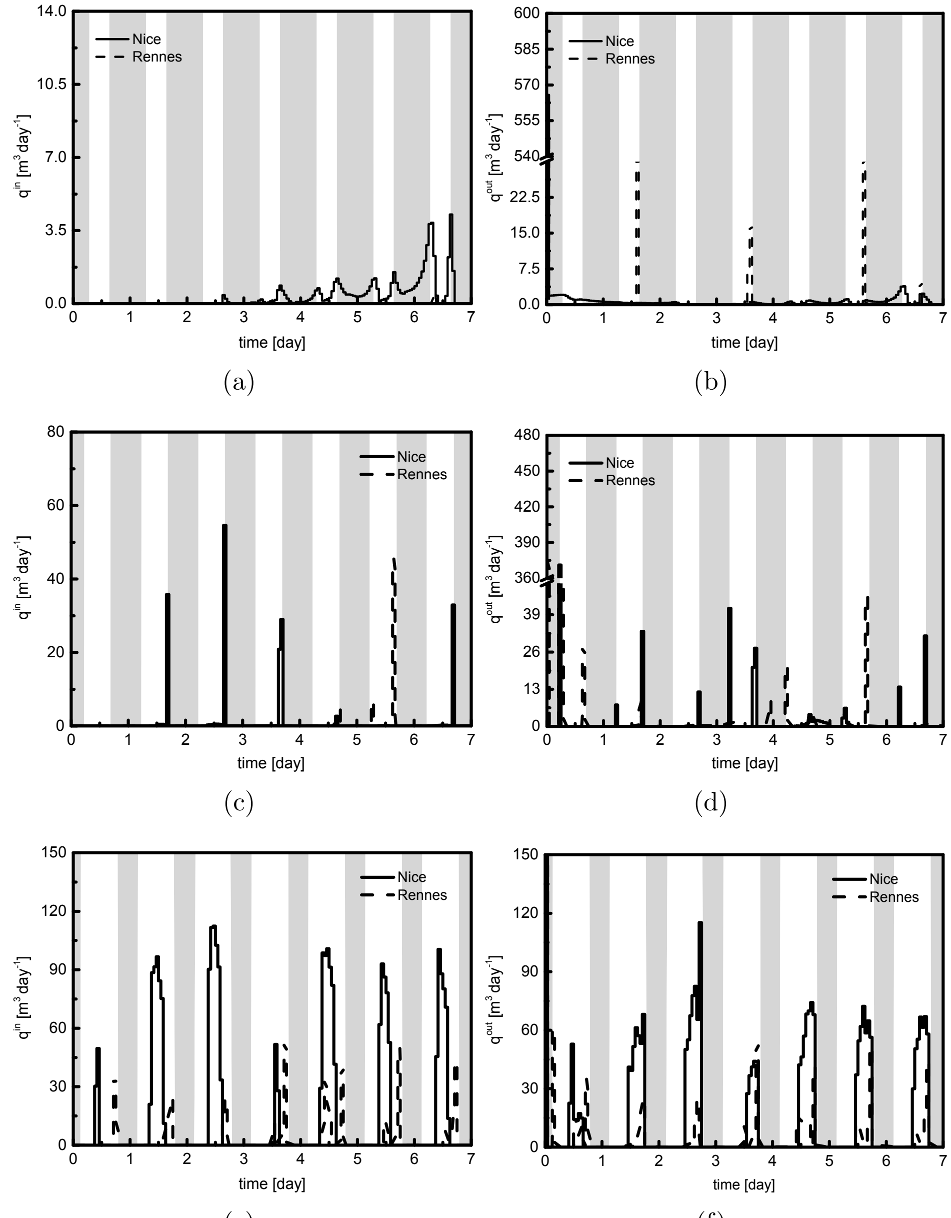
Case studies		Productivity (kg·week^{-1})		Water demand (m^3·week^{-1})	
		Nice	Rennes	Nice	Rennes
Winter	Standard cultivation strategy	2.59	1.25	17.53	17.53
	Optimal control strategy	3.73	2.12	26.83	25.02
Spring	Standard cultivation strategy	5.60	3.96	17.53	17.33
	Optimal control strategy	9.02	5.61	32.60	27.48
Summer	Standard cultivation strategy	11.62	12.71	17.53	17.53
	Optimal control strategy	25.83	19.59	122.98	45.97



(a) (b)

(c)

Figure 4.1: *Optimal weekly P_{net} dynamics obtained for Nice (solid line) and Rennes (dashed line) in: (a) Winter, (b) Spring, (c) Summer. (The background is colored in white at daytime and in grey at nighttime.)*



(a) (b)

(c) (d)

(e) (f)

Figure 4.2: *Optimal weekly q^{in} and q^{out} dynamics obtained for Nice (solid line) and Rennes (dashed line) in: (a, b) Winter, (c, d) Spring, (e, f) Summer. (The background is colored in white at daytime and in grey at nighttime.)*

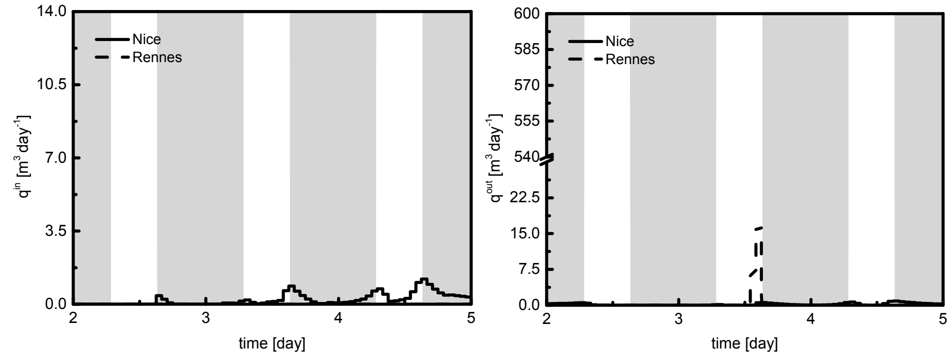
Interestingly, Table 4.1 shows that productivity was slightly higher in Rennes than in Nice in Summer under standard operation. This result is explained by the occurrence of high temperature peaks in Nice, which cause productivity to significantly drop. The importance of temperature on productivity is further discussed in the following discussion. Fig. 4.2 shows the optimal q^{in} and q^{out} profiles maximizing algal productivity over the entire cultivation period. Fig. 4.2 reveals that medium injection or culture extraction only occurred at day time under optimal operation. Although the resulting control strategy was different for the two locations, a qualitatively recurrent behavior can be identified despite the weather variability between different days of cultivation and the different periods of the year. The behavior of the optimizer was therefore analyzed on a time window of 3 cultivation days only.

4.2 Optimal operation strategy: key features

Before analyzing the control strategy followed by the optimizer, a brief summary of the ideal control logic ensuring the highest possible productivity is described in this paragraph. Firstly, algal concentration must be optimized at daytime by accounting for two processes: 1) biomass losses through respiration linearly increase with the algal concentration, and 2) the amount of light intensity captured by algal cells, hence photosynthetic rate, increases with algal concentration. As a result, there is an optimal algal concentration that should ensure that most of the light entering the pond is captured by algae while still maintaining respiration rates at a low value. Previous studies show that this optimal concentration is reached when the specific rate of photosynthesis at the pond bottom equals the specific rate of respiration (Takache et al. (2010)). Mathematically, these conditions are reached when the 'compensation function' f_{comp} defined in the previous Chapter is equal to 1. In summary, the ideal optimal biomass concentration at daytime $x_{b,opt}(t)$ is the algal concentration that guarantees that the compensation function defined in Eq. 3.36 equals 1. Secondly, maximal productivity is achieved when the pond temperature $T_p(t)$ is maintained at T_{opt} at daytime. At nighttime, the pond temperature $T_p(t)$ and the biomass concentration x_b(t) values should be maintained as low as possible in order to minimize respiration rates, hence biomass losses. The ideal optimal pond operation would therefore require a drastic change of the algal concentration and pond temperature at sunrise and sunset to ensure that these two variables are maintained at their optimal values at daytime and nighttime. Such drastic changes are in practice very difficult to achieve and the next paragraph discusses how the optimization scheme investigated in this Thesis handled these practical difficulties.

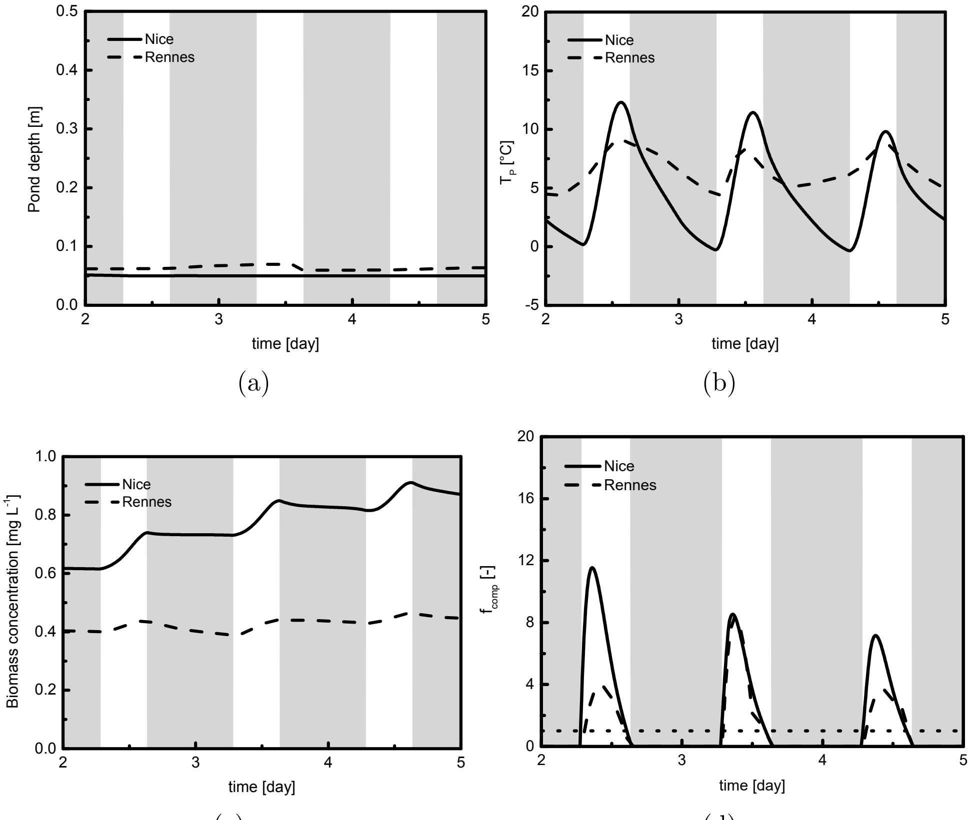
4.3 Detailed analysis of the optimization scheme

The analysis of the optimization scheme is split into four phases, from morning to night. *Morning.* Focusing first on the Summer case study, Figures 4.7(a) and 4.7(b) show that no water was injected to or extracted from the pond in the morning (q^{in}, $q^{out} = 0$), which led to maintain the pond depth in Rennes at a constant and low value ($l_p(t) = 0.05$ m, see Fig. 4.8(a)). Very small depths indeed minimize the thermal inertia of the pond and thus allow a fast increase of the pond temperature $T_p(t)$ (see Fig. 4.8(b)), hence a greater productivity increase. The same control strategy was used in Nice although the pond depth in Nice was slightly above the minimal physical constraint of 0.05 m (see Fig. 4.8(a)). This result can be explained by the fact that removing culture from the pond would lead to lower the biomass content and therefore increase the compensation function. Yet, the compensation function is already significantly higher than 1 in the morning in Nice (Fig. 4.8(d)), which means that removing more biomass would cause productivity losses. In Spring and Winter the morning control strategy was similar to Summer (Fig. 4.3(a), Fig. 4.3(b), Fig. 4.5(a) and Fig. 4.5(b)), if we neglect very small control peaks used to injected the culture medium when



(a) (b)

Figure 4.3: *Three-days zoom of: (a) q^{in} and (b) q^{out} dynamics obtained for Nice (solid line) and Rennes (dashed line) in January. (The background is colored in white at daytime and in grey at nighttime.)*



(a) (b)

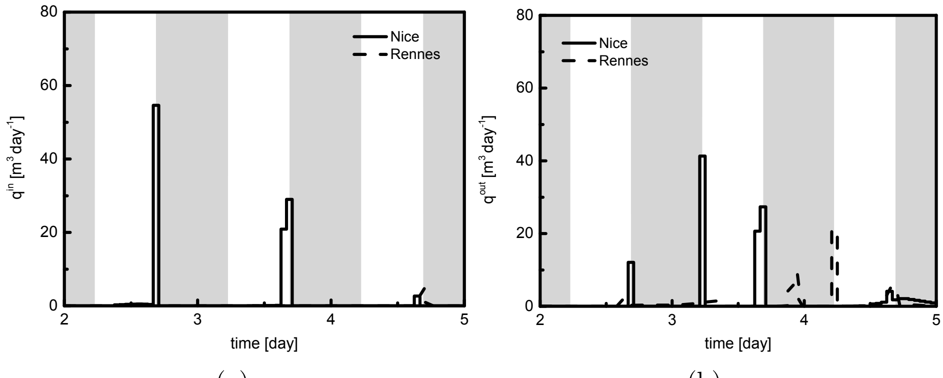
(c) (d)

Figure 4.4: *Three-days zoom of: (a) l_p, (b) T_p, (c) x_b and (d) f_{comp} dynamics obtained for Nice (solid line) and Rennes (dashed line) in January. (The background is colored in white at daytime and in grey at nighttime.)*

T_{in} was hotter than the algal culture in the pond at sunrise. Based on these observations, the optimizer behavior in the *Morning* phase can be schematized by the following simple rules:

- During the morning the pond depth is maintained as low as possible in order to help reaching both optimal pond temperature and biomass concentration as soon as possible at the start of the afternoon;
- In winter, if pond temperature is lower than inflow temperature and if the biomass content in the pond is high enough to avoid 'washout' conditions, the culture can be partially replaced with fresh medium to increase the pond temperature.

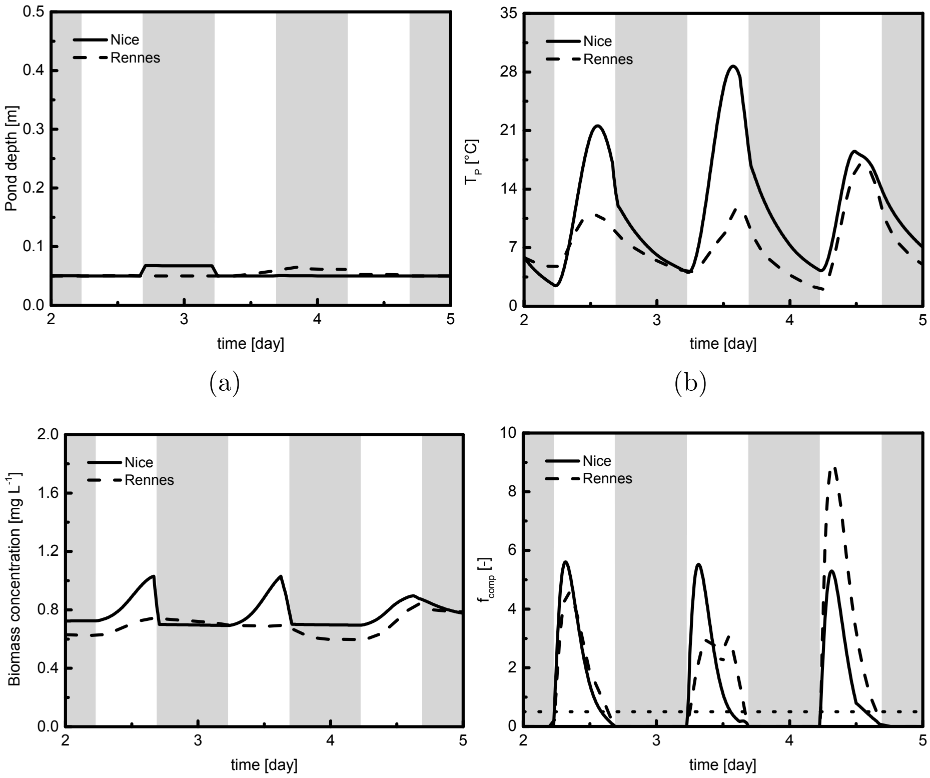
Afternoon. Figures 4.7(a) and 4.7(b) show that the inflow rate q^{in} in Nice case exhibits a 'bell curve' profile from midday until late afternoon in Summer. q^{out} followed the same dynamics but started slightly later in the day. In other words, the control strategy was mainly based on replacing the pond culture by fresh medium ('flushing' strategy). This culture replacement had mainly two consequences. Firstly, as shown in Fig. 4.8(d), the compensation function f_{comp} was maintained at a value close to 1 during the afternoon, indicating that the algal concentration was at its optimal value $x_{b,opt}(t)$ during the afternoon. Secondly, replacing algal culture by relatively cold fresh medium helped maintaining pond temperature close to its optimal level T_{opt} (35.8 °C, Fig. 4.8(b)). Fig. 4.8(a) shows that the pond depth l_p in Nice increased until mid-afternoon and then decreased, which indicates that culture replacement was not sufficient to maintain pond temperature at the optimal level.



(a) (b)

Figure 4.5: *Three-days zoom of: (a) q^{in} and (b) q^{out} dynamics obtained for Nice (solid line) and Rennes (dashed line) in March. (The background is colored in white at daytime and in grey at nighttime.)*

Increasing the pond depth indeed limited high temperature increases by increasing the pond thermal inertia. In summary, temperature control is the result of two combined strategies: culture replacement ('flushing strategy') by cold fresh medium and thermal increase through depth increase ('depth strategy'). Figures 4.7(a), 4.7(b) and 4.8(a) show that the same 'flushing' and 'depth' strategies were also used in Rennes in Summer during



(a) (b)

(c) (d)

Figure 4.6: *Three-days zoom of: (a) l_p, (b) T_p, (c) x_b and (d) f_{comp} dynamics obtained for Nice (solid line) and Rennes (dashed line) in March. (The background is colored in white at daytime and in grey at nighttime.)*

day 5, but not during days 3 and 4. As Fig. 4.8(d) shows that the compensation function was significantly lower than 1 during the afternoons of days 3 and 4 in Rennes, further culture replacement could have theoretically been used to maintain the biomass concentration at its optimal value. Yet, replacing the culture by cold fresh medium (T_{in} was set at 16.5 °C in Rennes) at a faster rate would have significantly decreased the pond temperature and therefore lower biomass productivity. Fig. 4.8(b) shows indeed that days 3 and 4 were relatively cold, differently from day 5 in which the pond temperature reached its optimal value. In other words, the optimizer found the best trade-off between optimal biomass concentration and optimal temperature conditions in the case of warm but not hot weather conditions. In addition, Fig. 4.8(a) shows that in Rennes the depth was maintained at its lowest value in the afternoon of days 3 and 4 (warm days) in order to maximize the temperature increase at daytime. Figures 4.3(a), 4.3(b), 4.5(a) and 4.5(b) show that both q^{in} and q^{out} were maintained at 0 in Winter and Spring in Rennes and Nice. As a result, the biomass concentration slightly increased at daytime (Figures 4.4(c) and 4.6(c)). In addition, the depth was left at its lowest value (0.05 m) all day long. These observations indicate that the optimal strategy during cold days consists of maintaining the pond temperature as high as possible during daytime even if biomass concentration is significantly higher than the optimal concentration. This is mostly due to low pond temperatures that ensured low biomass losses through respiration during Spring and Winter. The optimizer behavior in the *Afternoon* phase can be schematized by the following simple rule:

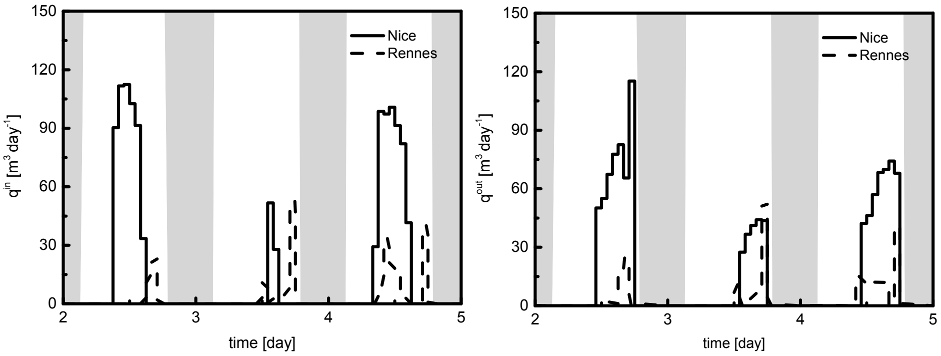
- In the afternoon the culture can be flushed to maintain the algal concentration at its optimal level. In Summer, this ‘flushing strategy’ can be combined with depth increase strategy to control temperature at its optimal level. In Spring and Winter, the optimal strategy consists of neither injecting fresh medium nor extracting culture during daytime while maintaining the pond depth at a low value, to ensure that pond temperature reaches the highest possible value.

Sunset. In Summer, Figures 4.7(a) and 4.7(b) show that a high fraction of the culture was replaced by fresh medium at sunset in Rennes. ‘Flushing’ the system at sunset both lowered pond temperature (see Fig. 4.8(b)) and biomass concentration (see Fig. 4.8(c)), which in return limited respiration rates at nighttime. The alternative strategy used in Nice was based on decreasing the pond depth when approaching sunset (Fig. 4.8(a)), which removed a significant fraction of the biomass from the pond. In addition, decreasing pond depth accelerates the pond temperature decrease at night. Similarly to the ‘flushing’ strategy used in Rennes, this ‘depth-decrease’ control strategy reduced respiration at nighttime. Removing all the biomass from the system would obviously ensure no respiration at nighttime but this would also cause the productivity to be null the day after. As a result, the optimizer finds the optimal algal concentration ensuring both low respiration rates at nighttime and high productivities the following morning. In Winter only a small fraction of the culture was replaced by fresh medium in Nice (Figures 4.3(a) and 4.3(b)) as night-time respiration rates were limited by cold temperatures (Fig. 4.8(b)). In Rennes the ‘flushing strategy’ at sunset was not applied, mostly because inflow temperature was higher (9.2 °C) than pond temperature $T_p(t)$ at sunset. Injecting relatively warm water at sunset would therefore only lead to high respiration rates at nighttime. The optimal strategy at Nice in Spring consisted on partly ‘flushing’ the system at sunset similarly to the summer case. In Rennes, no culture was replaced at sunset in Spring, mostly because maintaining temperatures as high as possible was the best strategy to optimize productivity (Figures 4.5(a) and 4.5(b)). The optimizer behavior at *Sunset* can be schematized by the following simple rule:

- During hot days, a fraction of the culture is replaced with fresh medium at sunset to minimize nighttime respiration rates. Pond depth is also maintained at a low level to ensure low nighttime temperatures. In Winter or in colder climates, no culture is replaced by fresh medium at sunset to ensure that temperature is as high as possible during the following day.

Night. The pond depth was maintained at its sunset value all night long independently on the season considered mostly (Figures 4.4(a), 4.6(a) and 4.8(a)). In addition, Figures 4.3(a), 4.3(b), 4.5(a), 4.5(b), 4.7(a) and 4.7(b) show that, in general, no ‘flushing’ was used at nighttime ($q^{in} = 0$ and $q^{out} = 0$). Some exception were reported in Spring at day 4 for Rennes and at day 5 for Nice (the peaks correspond to culture extraction just after rainfall to maintain the pond depth at its lowest possible value) or in Winter at day 4 and 5 in Nice, where slight culture replacements were done both to decrease the respiration rates by acting on biomass reduction and to avoid too low temperatures for the following day of cultivation. In Summer, this strategy ensured low temperatures at nighttime. In Winter, maintaining a low depth at daytime is necessary to reach high productivity values and practically constrains low pond depths at nighttime. The optimizer behavior at *Night* can be schematized by the following simple rule:

- The depth is maintained at the value set at sunset and no water is injected to or extracted from the pond.

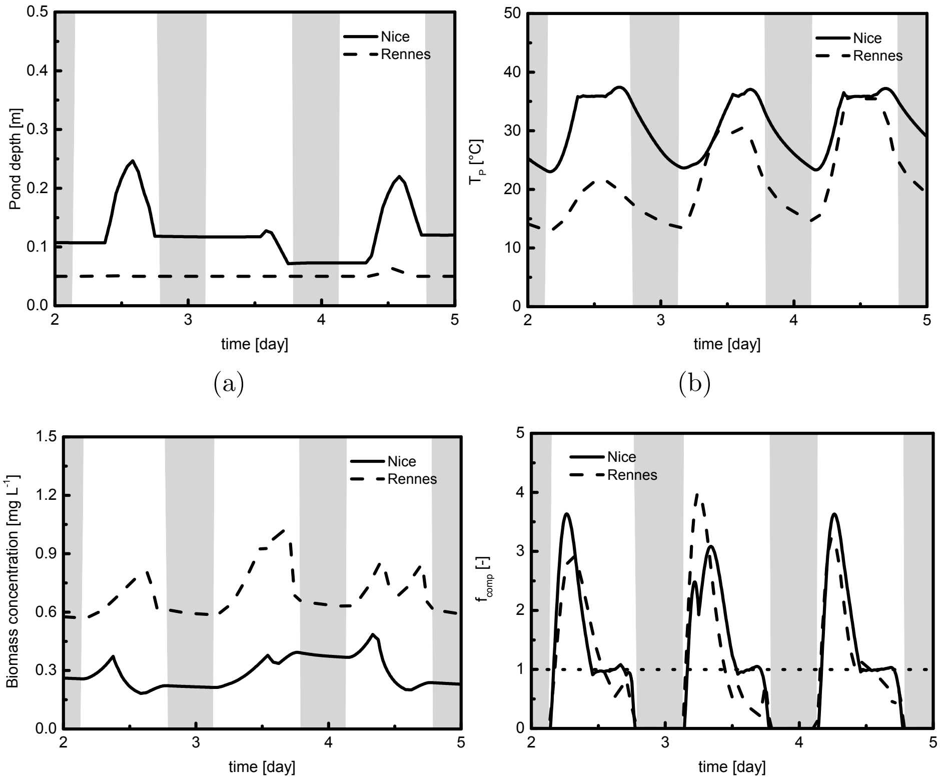


(a) (b)

Figure 4.7: *Three-days zoom of: (a) q^{in} and (b) q^{out} dynamics obtained for Nice (solid line) and Rennes (dashed line) in July. (The background is colored in white at daytime and in grey at nighttime.)*

4.4 Discussion

The optimization technique significantly increased the productivity at the two locations and three seasons considered. The productivity boost in summer mainly results from the optimizer ability to maintain, during large periods of daytime, ideal growth conditions, *i.e.* optimal concentration and temperature (via the ‘flushing’ and ‘depth’ strategies). In Spring and Winter, the optimal temperature for the species *Chlorella vulgaris* (35.8 °C) cannot be maintained, so the optimal strategy consists of limiting culture replacement to ensure high pond temperatures, even if this strategy leads to relatively high biomass concentrations. Because of relatively low temperatures, respiration rates are indeed relatively low, so these high biomass concentrations do not lead to important biomass losses. The knowledge of future weather is crucial to optimize the fresh medium injection rate and culture extraction rate and this can be illustrated in several cases. Firstly, in hot days, slowly increasing the pond depth can help to maintain the pond temperature at its optimal value during daytime. As temperature change of water bodies is a relatively slow process due to the high thermal inertia of water, only an accurate knowledge of future weather conditions and their impact on pond temperature can lead to a control of pond temperature at its optimal value. Secondly, a fraction of the algal culture is replaced by fresh medium at sunset to minimize respiration losses at nighttime. However, removing too much biomass from the pond would lead to low productivity values the morning after, and especially if the day after is particularly sunny. Determining the optimal fraction of culture to remove from the pond at sunset therefore requires knowing the weather conditions of the following day. Temperature control by playing on thermal inertia and culture replacement by fresh medium is central in the optimization strategy. This point was so far never considered in the previous optimization studies which focused on optimizing algal concentration through adjustment of the compensation function. *Chlorella vulgaris* is relatively resistant to high temperatures ($T_{max} = 42$ °C). However, for a cold-adapted algae species, the possibility of culture crashes due to relatively high temperatures (*i.e.* above T_{max}) would increase further the impact of temperature on productivity. In this case, the optimizer would likely place temperature control above concentration control in summer to avoid culture crashes.



(a)

(b)

(c)

(d)

Figure 4.8: *Three-days zoom of: (a)* l_p*, (b)* T_p*, (c)* x_b *and (d)* f_{comp} *dynamics obtained for Nice (solid line) and Rennes (dashed line) in July. (The background is colored in white at daytime and in grey at nighttime.)*

In this study, we did not include the water cost in the optimization criterion. However, the optimal management induces high dilution rates, especially in Summer, both to reduce temperature and to dilute culture at high density. Water demands were consequently relatively high (Table 4.1) as the standard management procedure always required a significantly lower amount of water (only 17.53 m^3 week^{-1} for all the seasons) than the optimal control strategy. In particular, the optimized control strategy increased the WD at Nice in Summer up to a factor 7, which means an increase by a factor 3.2 per kg of produced algae: from 1.5 m^3 kg^{-1} up to 4.76 m^3 kg^{-1}. This computation also highlights the necessity to recycle water after biomass extraction to reduce the water need. As a perspective, the same optimization strategy should be applied but by constraining the amount of water that can be used in the process (based for example on the availability on rainwater at the location considered), and assuming that an important fraction of the water can be recycled, as suggested by White and Ryan (2015).

Table 4.2: *Optimal daily management of ideal ponds: the ideal strategy*

Key variables	Night	Day
T_p	as low as possible	as near as possible to T_{opt}
x_b	as low as possible	as soon as possible near to $x_{b,opt}(t)\|f_{comp}(t) = 1$

4.5 Optimization approach for MPC implementation

Model Predictive Control (MPC) (Camacho and Alba (2013)) has proven to be very efficient to manage situations with complex modeling, where classical model based control cannot be easily derived. MPC seems especially relevant when meteorology plays a key role, such as energy efficient building climate control (Oldewurtel et al. (2012)), or management of distributed power system with wind turbines (Zong et al. (2012)). MPC was already applied to microalgae (Berenguel et al. (2004); Tebbani et al. (2014a,b)), but never to manage weather forecasts in the control strategy. Alternatively, here we propose an optimization strategy to be used in a MPC framework, accounting for the knowledge of future weather conditions. We first detail the numerical approach developed for solving the problem, then we consider the case where weather forecasts are perfect and we analyze the logic behind the control action. Finally, we consider the more realistic case where meteorology becomes uncertain after 24 hours. We show that our approach maintains the productivity increase by a factor 2.2 compared to the reference case of constant dilution rate and raceway pond depth.

4.5.1 Objective criterion

The control target was the maximization of microalgal productivity on a moving time window from the current time τ_n to τ_n+T_h. It consists in finding the optimal fresh medium volumetric flow rate (q^{in}) and the culture volumetric extraction rate (q^{out}). The control vector $\mathbf{q} = (q^{in},$

Control variables switching frequency	Control horizon		
	1 day	2 days	4 days
2 hours	T_r	1.98 T_r	4.04 T_r
1 hour	1.32 T_r	4.29 T_r	11.47 T_r
1/2 hour	8.17 T_r	33.64 T_r	209.19 T_r

Table 4.3: *Computational cost evaluation for different discretization times and horizons. The reference time T_r for a Samsung Electronics laptop, Intel(R) Core(TM) i7-3635QM CPU @ 2.40 GHz, 8.00 GB RAM, was T_r=13.3 minutes.*

$q^{out})^{\mathsf{T}}$ was therefore the solution of the optimal control problem over the time horizon τ_n+T_h:

$$\begin{array}{ll} \max_{\mathbf{q}} P_{net} = & \max_{\mathbf{q}} \int_{\tau_n}^{\tau_n+T_h} (G(x_b, H_s, T_p) - R(x_b, T_p))V dt \\ \text{s.t.} & \\ & 0 \le q_i(t) \le q_{\max} \\ & l_{p,\min} \le l_p(t) \le l_{p,\max} \\ & \dot{\boldsymbol{\xi}} = \mathbf{g}(\boldsymbol{\xi}, \mathbf{q}, t), \quad \boldsymbol{\xi}(0) = \boldsymbol{\xi}_0 \end{array} , \tag{4.1}$$

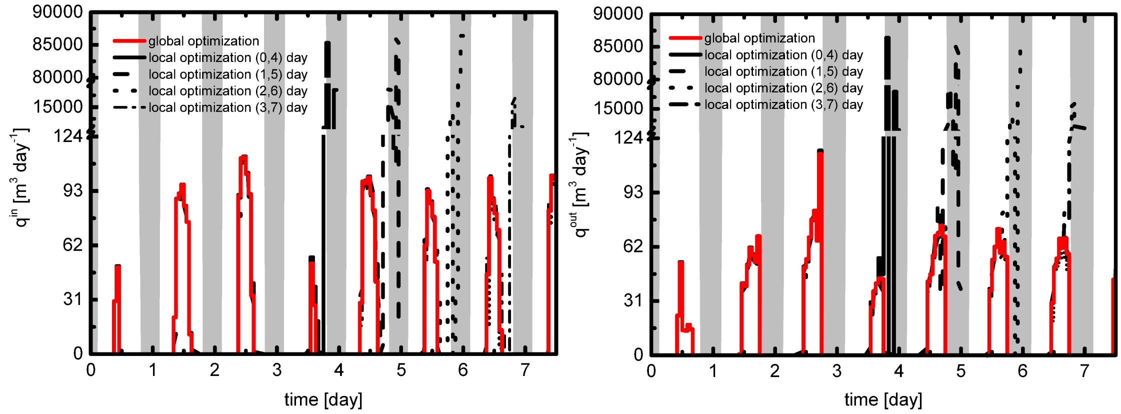
where $\boldsymbol{\xi}$ is the state vector and $q_{\max}$ is the upper bound of the flow rates. The dynamical system $\mathbf{g}(\boldsymbol{\xi}, \mathbf{q}, t)$ therefore gathered the biological and thermal dynamics (see Equations (3.1) and (3.9)). The upper bound of the two flow rates is $q_{\max}$. The objective function P_{net} is the integral with respect to time of the difference between growth and respiration during the considered time window, therefore representing the net algal productivity.

The computed optimal control is then applied for the period (τ_n, τ_{n+1}). Since new weather forecasts are available at time τ_{n+1}, a new control is then computed for the new period (τ_{n+1}, $\tau_{n+1} + T_h$). The procedure was iterated until the end of the cultivation period (at day 7). We assume that temperature and biomass density are perfectly on-line measured and known at the beginning of each new optimization period (biomass can *e.g.* be assessed by simple measurements of the optical density Havlik et al. (2013)). The optimization task was implemented through gPROMS$^{\text{TM}}$ (4.1 version) default solver NLPSQP, which uses a sequential quadratic programming SQP method to solve nonlinear programming NLP problems (optimization tolerance = 0.001). The two control inputs q^{in} and q^{out} were numerically implemented as piecewise constant within the range [0-1] m^3/s. Each new optimization problem is initialized with the flow vector resulting from the previous optimization which both reduced the computation time and decreases the risk of local minima.

Table 4.3 shows the computational times required for different optimization tasks. The reported results were obtained through the variation of the switching frequency of each input variable (30 mins, 1 hour, 2 hours) and the implementation of different control horizons (1 day, 2 days and 4 days). We selected a reasonable trade-off between accuracy and computation time with a 1 hour input discretization time for a 4 day horizon. This required 2.5 hours on our computer, which is realistic for an on-line implementation. Each case study was initialized by setting the two control variables (q^{in} and q^{out}) equal to 10^{-4} m^3 s^{-1}.

4.5.2 Comparison with a reference management strategy

This first study was based on a perfect weather forecast. In Figures 4.9(a) and 4.9(b) the optimal controlled variables profiles obtained by considering both the global optimization (τ_0, T_f) and the sequence of sub-optimization tasks defined in the previous paragraph (with T_h set equal to 4 days) are reported. No significant variations between the two approaches were registered, except for the final peaks of q^{in} and q^{out} at the end of the final day of each interval (τ_n, $\tau_n + T_h$). Such behavior is related to the fact that the 4th day of each time window is considered by the optimizer as the last day of cultivation. A complete emptying of the open raceway pond was therefore required to collect all the residual biomass in the system. However, since the optimization update was carried out every day, no practical effect of such peaks on the raceway operation was registered and almost a perfect superimposition of the global and local optimization curves is obtained. In the ideal case, the optimal control problem should be solved on the interval (τ_n, T_f). However, the dimension of this problem when T_f - τ_n is larger than 5 days induces a high computational cost, and the risk of non convergence due to non convexity. Of course, the larger T_h the better, but the choice of T_h is practically determined by the computation time to solve the problem on a larger interval. It turns out that (see Figures 4.9(a) and 4.9(b)) $T_h = 4$ days leads to an efficient trade-off, while considering a larger time window did hardly influence the optimal solution on the interval (τ_n, τ_{n+1}).



(a) (b)

Figure 4.9: *Global and local optimal profiles of: (a) q^{in} and (b) q^{out}. Global optimization (red line) represents the profiles obtained by taking into account the time horizon (τ_0, T_f) whereas the local optimal profiles (black lines) represent the results obtained for all the sub-optimization tasks conducted in the intervals (τ_n, $\tau_n + T_h$), with n=0:6 and T_h=4. (c) shows the profiles obtained with the global optimization (solid line) and the standard control of open ponds (Jorquera et al. (2010), Rogers et al. (2014)) (dashed line). The background is colored in white at daytime and in grey at nighttime.*

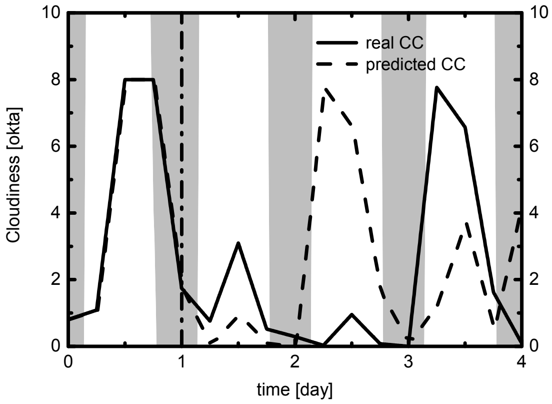


Figure 4.10: *The real and predicted profiles for cloudiness are reported. The solid line represents the real weather dynamics whereas the dashed line represents the weather forecast used for the optimization. (The vertical dashed-dotted line represents the artificial division between unaltered and modified weather data for a generic* (τ_n, $\tau_n + T_h$) *time window.)*

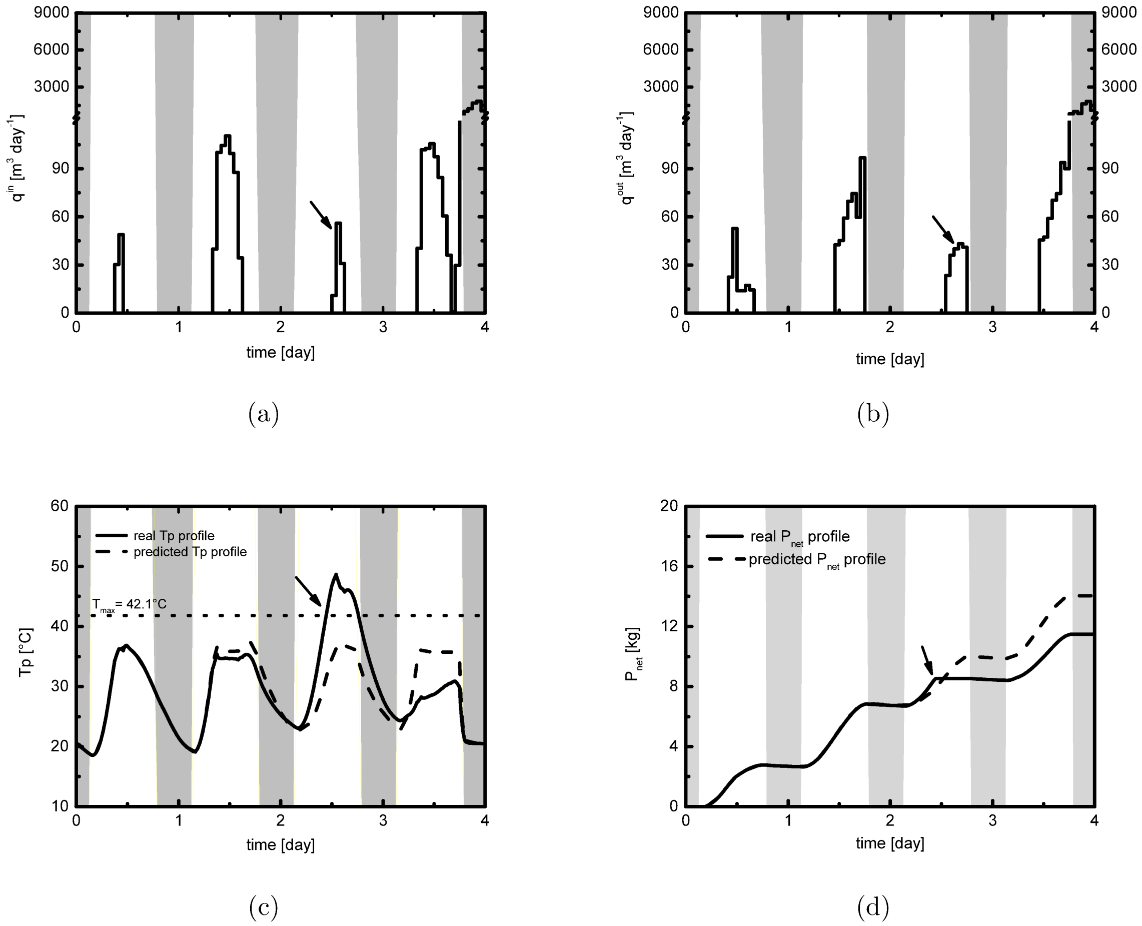
4.5.3 Numerical implementation of weather forecast uncertainty

Despite the increase in meteorological accuracy and computing power, the prediction capability is in the range of a few days. To simulate weather forecast uncertainty, we assumed a perfect forecast during the first day and we simply shifted the meteorological data by one day for the rest of the period. This is probably a strong perturbation, in comparison with the accuracy of the existing meteorological models (see Lorenz et al. (2009b); Perez et al. (2013) for a detailed study) but it generates a credible weather forecast. This is illustrated on Fig. 4.10 for cloudiness data, but the logic was used for all the five weather variables extracted by ECMWF website. More accurately, for each preset optimization time window (τ_n, $\tau_n + T_h$) (where τ_n is the current time and T_h is the length of the considered time horizon) the weather forecast data was corrupted in the following way:

- No discrepancies between real and predicted data were introduced in the time window (τ_n, τ_{n+1})
- For each j-th day in the interval (τ_{n+1}, $\tau_n + T_h$), the data used for the optimization were the same collected for the ($j + 1$)-th day; as an example, if we consider the first 4 days of cultivation, the data used for day 2 were the same as the real ones collected for day 3, and so on till day 4

4.5.4 Impact of inaccurate weather on the optimal solution over the (τ_n, $\tau_n + T_h$) time window

In this paragraph we assess the impact of inaccurate weather forecasts. Optimizing the control strategy by setting large time windows and, consequently, using highly uncertain weather forecasts, has a great impact on the net productivity of the system. Fig. 4.11(d)



(a) (b)

(c) (d)

Figure 4.11: *(a) and (b) show the optimal control values of q^{in} and q^{out} by taking into account the uncertain weather forecast data reported in Fig. 4.10, whereas in (c) and (d) the effect of forecast uncertainty on T_p and net productivity is represented . (The background is colored in white at daytime and in grey at nighttime. The arrows highlight the negative effects of wrong forecasts on net productivity as a consequence of q^{in} and q^{out} optimal profiles.)*

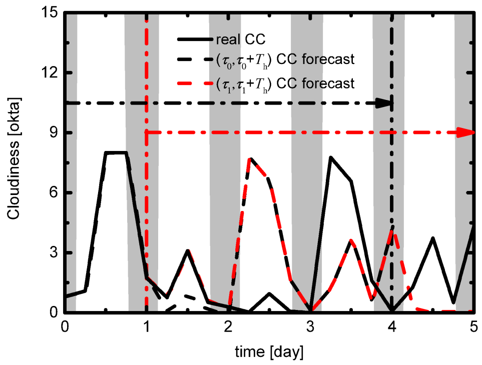


Figure 4.12: *Five-days zoom of the CC profiles. The solid black line represents the real weather data, whereas the black and red dashed lines represent, respectively, the weather forecast for the first and the second optimization windows. The horizontal arrows and the dashed-dotted vertical lines are used to point out each optimization window. The first window (τ_0, τ_0 + T_h) is in black and the second (τ_1, τ_1 + T_h) is in red (the background is colored in white at daytime and in grey at nighttime.)*

shows that at day 3 there is a loss of productivity with respect to the predicted profile. This critical point affected the productivity of the 4-day time window by lowering its value of 3.2 kg with respect to the ideal case previously described. This result can be explained by considering the q^{in} and q^{out} optimal profiles. Focusing on day 3, Figures 4.11(a) and 4.11(b) show low control peaks due to the fact that a cloudy day was wrongly predicted. Since day 3 was a sunny day the optimal control strategy was not sufficient to cope with high temperatures and high solar irradiances. In fact, Figure 4.11(c) shows that at day 3 the raceway pond temperature overcame the T_{max} value hindering the growth rate of the culture.

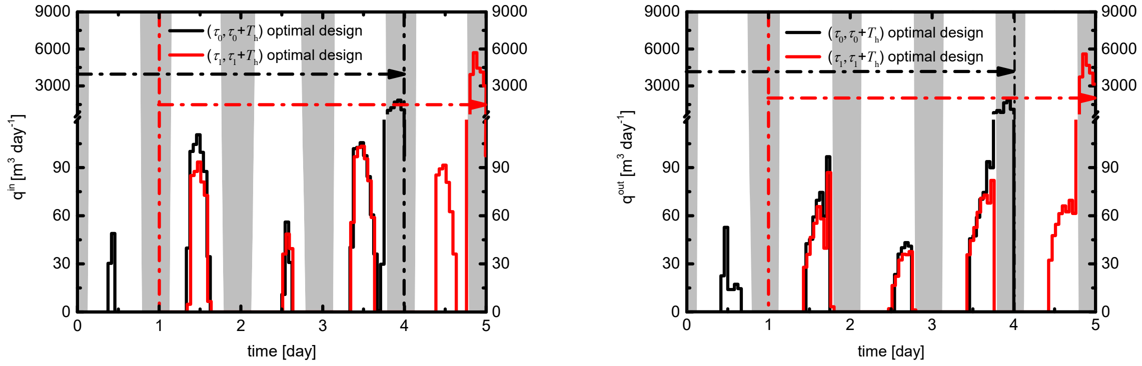
Now, in the next paragraph, we examine how a daily update of the weather forecasts together with a 4 day time window can compensate for the possible impact of a wrong meteorology.

4.5.5 Optimization with inaccurate meteorological forecasts

In this section, the predictive controller is implemented with $T_h = 4$ days by using the approach described in section 4.5.

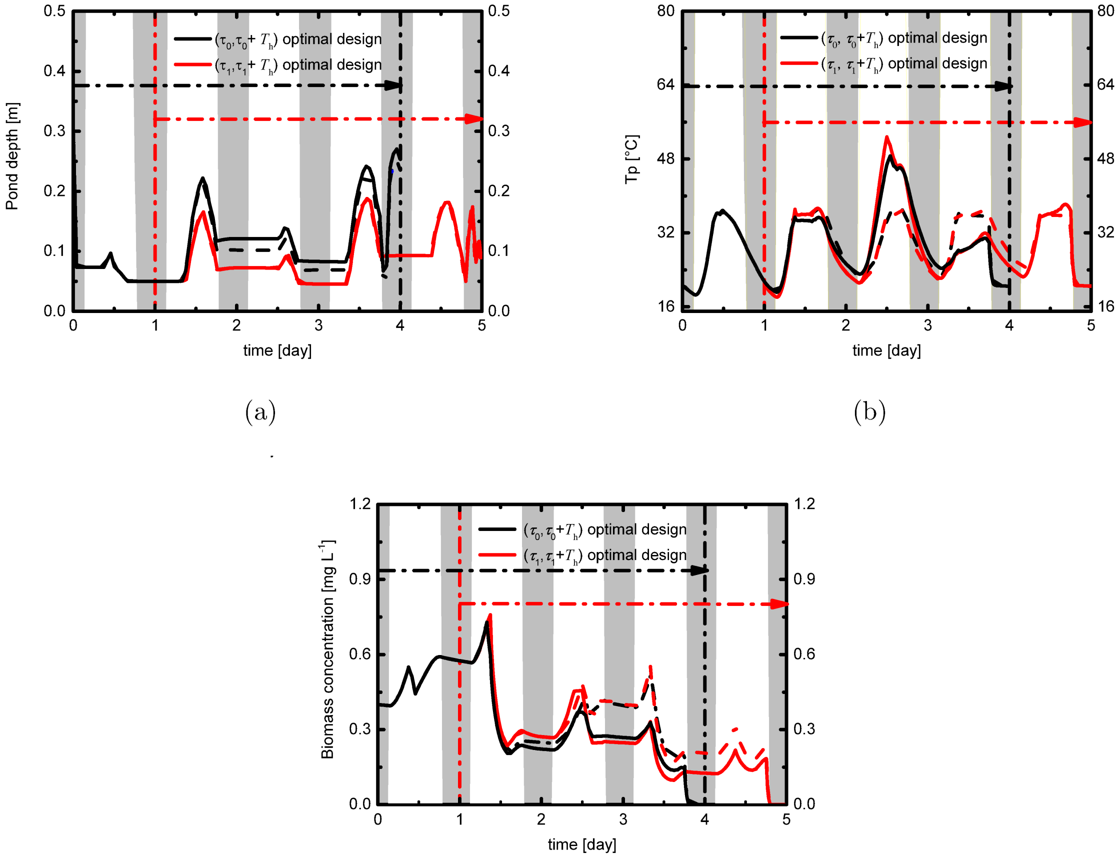
For the sake of simplicity the results are reported only for the first two optimization iterations. The inflow rate q^{in} and the outflow rate q^{out} profiles are reported in Figures 4.13(a) and 4.13(b) whereas the raceway pond depth $l_p(t)$, the raceway pond temperature T_p(t) and the biomass concentration $x_b(t)$ are reported in Fig. 4.14.

This optimization approach with daily re-optimization allows to maintain productivity values close to the ideal case (25.83 kg). These results demonstrated that the optimal control strategy is mainly linked to weather conditions related to the first days of optimization and that a daily update seems enough to prevent the negative effect of weather forecast uncertainty.



(a) (b)

Figure 4.13: *(a) q^{in} and (b) q^{out} optimal control values. The horizontal arrows and the dashed-dotted vertical lines are used to point out each optimization window. For the sake of simplicity, only the first two optimization windows (τ_0, $\tau_0 + T_h$) (in black) and (τ_1, $\tau_1 + T_h$) (in red) are represented. (The background is colored in white at daytime and in grey at nighttime.)*



(a)

(b)

(c)

Figure 4.14: *(a) l_p, (b) $T_p(t)$ and (c) $x_b(t)$ optimal control values. The horizontal arrows and the dashed-dotted vertical lines are used to point out each optimization window. For the sake of simplicity, only the first two optimization windows (τ_0, $\tau_0 + T_h$) (black) and (τ_1, $\tau_1 + T_h$) (red) are represented: solid lines represent the real system behavior whereas dashed lines represent the predicted dynamics. (The background is colored in white at daytime and in grey at nighttime.)*

4.6 Conclusions

The main result of our study is the key role played by temperature control for optimizing microalgal productivity. Several studies have focused on control strategies deriving an optimal microalgal density, balancing growth and respiration at the bottom of the reactor (compensation condition). The control strategy however gives to this target only the second rank. It first adapts the raceway pond volume and flow rates in order to get as close as possible to the optimal temperature for growth, and always prevent the temperature to go beyond this optimal value. The analysis of the rational behind the optimal control revealed that this productivity gain was achieved via two main mechanisms: 'flushing' the culture and controlling the raceway pond depth. The rather periodical control actions give the pace for inflow and outflow. The compensation condition is then tracked once temperature is close to the optimal value. This control scheme turns out to significantly increase the productivity of algal open raceway ponds compared to standard operation with constant depth and dilution rates. The proposed control strategy is highly dependent on the weather forecasts, and an accurate prediction must be available at least for one day to guarantee an efficient process optimization. Inaccurate predictions can be compensated by a more frequent optimization update. It is shown that the complex optimal strategy can be reduced to six simple rules whose implementations are simpler. Using these simple rules to optimize raceway operation has therefore the potential to: (*i.*) minimize computational costs, (*ii.*) supply guidelines for practical operation of open raceway ponds and (*iii.*) develop simple closed loop control loops in order to reduce the sensitivity of the control strategy to inaccurate weather forecasts. It should be noted that the proposed model does not take into account the effect of important mechanisms, such as photoacclimation. However, the frequent culture replacements would probably reduce photoacclimation impact on final microalgae productivity, but further work is needed to validate this statement. Secondly, the optimization reliability in terms of reproducibility should be further checked: some preliminary tests were conducted in order to compare the optimal profiles obtained with different initializations, showing negligible discrepancies between the different cases. Nonetheless, advanced optimization techniques (e.g. multiple shooting) could be used in order to avoid local optima/sub-optimal solutions. Finally, a post-optimal sensitivity analysis could be implemented in order to establish an updated confidence interval for the most impactful model parameters.

Chapter 5

Feasibility approaches for open pond management with inaccurate weather forecasts

This Chapter proposes a feasibility analysis to prevent both high productivity loss due to wrong control management and potential critical conditions caused by inaccurate weather predictions (e.g. cell death due to high temperatures). In particular, the Chapter was based on the evaluation and comparison of: (*i*) a constrained productivity optimization based on the implementation of a dynamic threshold (Kookos and Perkins (2004)) generated by the actual error on weather data measurements; (*ii*) a conservative control strategy generated through the iterative implementation of the daily 'worst case' meteorological scenario, regardless of the forecast data extracted. The objective of this study is to investigate how these two approaches can guarantee both high productivity levels and feasible operation in case of inaccurate weather forecasts. The secondary objective is to discriminate between the two approaches in terms of productivity loss and water demand increase, with respect to the ideal case of perfect weather predictions. The third objective is to implement the open pond in a typical biofuels production process in order to evaluate the effects of each proposed approach on total process cost.

5.1 Description of weather data inaccuracy

The representation of inaccurate weather data was based on modifying the real data extracted by the European Centre for Medium-Range Weather Forecast ECMWF website. Firstly, we conducted a preliminary sensitivity analysis in order to evaluate the most impactful weather variables in terms of productivity loss. The sensitivity analysis on productivity was done through the implementation of the same relative error (10%, -10%, 50% or -50%) to each meteorological data. It should be precised that the relative error on T_a was calculated with respect to Celsius values. Moreover, a physical threshold was assumed as the maximum allowable value for both CC and RH (8 okta and 100%, respectively).

The sensitivity analysis, conducted at different seasons in Nice location (Winter, Spring, Summer 2012, see Tables 5.1, 5.2, 5.3), demonstrated that CC and T_a are the most significant variables especially in the hottest and sunniest period of the year. In addition to this, in case of unpredicted high T_a and low CC, the actual control system could be unable to avoid critical conditions for cultivation (for example, T_p higher than T_{max}, with consequent microalge death).

Forecast variable	Deviation from real meteo data (%)			
	-50%	-10%	+10%	50%
Relative humidity RH	-4.85%	-0.93%	+0.90%	+3.61%
Cloudiness CC	+10.1%	+3.53%	-2.05%	-5.34%
Rainwater rate v_r	+0.55%	+0.11%	+0.11%	-0.54%
Air temperature T_a	-7.49%	-1.46%	+1.44%	+6.95%
Wind velocity v_w	+2.93%	+0.50%	-0.47%	-2.10%

Table 5.1: *Productivity variation in January (%) with respect to the optimal productivity obtained when real data coincides with predicted data.*

Forecast variable	Deviation from real meteo data (%)			
	-50%	-10%	+10%	50%
Relative humidity RH	-2.94%	-0.57%	+0.55%	+2.16%
Cloudiness CC	+17.53%	+6.01%	-2.13%	-4.99%
Rainwater rate v_r	+0.95%	+0.18%	-0.18%	-0.87%
Air temperature T_a	-4.41%	-0.86%	+0.85%	+4.16%
Wind velocity v_w	+6.30%	+1.09%	-0.97%	-3.96%

Table 5.2: *Productivity variation in March (%) with respect to the optimal productivity obtained when real data coincides with predicted data.*

Forecast variable	Deviation from real meteo data (%)			
	-50%	-10%	+10%	50%
Relative humidity RH	-0.56%	-0.09%	+0.068%	+0.15%
Cloudiness CC	-8.64%	+2.26%	-2.26%	-6.63%
Rainwater rate v_r	+0.28%	-0.01%	-0.01%	-0.21%
Air temperature T_a	-3.91%	-0.48%	+0.22%	-5.59%
Wind velocity v_w	-0.79%	+0.02%	-0.15%	-1.20%

Table 5.3: *Productivity variation in July (%) with respect to the optimal productivity obtained when real data coincides with predicted data.*

The inaccurate weather data was therefore represented through the addition of an artificial error term to each 6-hourly value of CC and T_a; conversely, the other three meteorological variables were kept constant at their real values.

The standard deviation of CC data extracted by ECMWF website was reported in the annual evaluation report of ECMWF (Haiden et al. (2015)). This report showed relative-mean-squared errors between 30-40% for cloudiness forecasts even on the first day of prediction. Nevertheless, the assumption of a universal constant error on cloudiness measurements neglected important aspects, such as the error variability related to the specific geographical conditions and the phase of the day at which each sample was collected. The works of Lorenz et al. (2009b), Perez et al. (2013) showed in fact that the relative mean square error of solar irradiance values (directly dependent on cloudiness data) given by ECMWF highly varies with both the forecast time horizon and the considered location (see Table 5.4). Furthermore, Lorenz et al. (2009a) proposed a correlation to directly estimate solar irradiance uncertainty through available ECMWF data. The correlation proposed by Lorenz et al. (2009a) was generated through experimental data collected at different locations in Germany. In this Thesis we used the same function correlation, but scaled up to a factor 2 (it corresponds to the addition of coefficient 2 (see Eq.5.1) to the original expression derived by Lorenz et al. (2009a)), in order to adapt it to the sunnier conditions in Nice during Summer (the highest solar irradiance peak (July 2012) in Nice was, in fact, twice as high as the maximum solar peak in Germany in the same season). Then, the standard deviation of solar irradiance measurements σ_{H_s} was defined as:

Table 5.4: *Relative mean square error of irradiance data collected by ECMWF website at different locations and different forecast time horizons (Lorenz et al. (2009b) and Perez et al. (2013))*

Location	1 day-horizon	2 days-horizon	3 days-horizon
Germany	40.3%	41.6%	44.9%
Switzerland	39.6%	41.8%	42.7%
Austria	45.6%	47.4%	50.5%
Spain	20.8% – 22%	21.3%–23%	22.4% – 23%
Central Europe	40%	42%	43%
Canada	32%	36%	-
USA	33%	35%	37%

$$\begin{aligned}\sigma_{H_s} &= 2 \times (8.455 - 138.6\cos\theta_z - 231kt^* + 1127\cos^2\theta_z + 123.5kt^*\cos\theta_z + \\ &+ 1315kt^{*2} - 2381\cos^3\theta_z + 290.8kt^*\cos^2\theta_z + 570.8kt^{*2}\cos\theta_z + \\ &- 2154kt^{*3} + 1583\cos^4\theta_z - 306.5kt^{*2}\cos^3\theta_z + 64.47kt^{*2}\cos^2\theta_z + \\ &- 735.2kt^{*3}\cos\theta_z + 1074kt^{*4}),\end{aligned} \tag{5.1}$$

where θ_z is the zenith angle (the angle between the vertical axis and the sun direction), whose calculation was described in Chapter 3. The variable kt^* is the clear-sky index, calculated through CC data with the following expression (Lorenz et al. (2009a)):

$$kt^* = \frac{4 - 3(CC/8)^{3.4}}{4}. \tag{5.2}$$

The standard deviation of T_a data extracted by ECMWF was assumed to be independent of local conditions and seasonal variations. The work presented by Haiden et al. (2015) shows that the average standard deviation of T_a is between 2 and 3 K. In this work the measurement error $\varepsilon_{T_a,i}$ associated to the i-th weather sample was defined as:

$$\varepsilon_{T_a,i} = \frac{\varepsilon_{T_a,i-1} + \varepsilon_{rnd,i}}{2}, \tag{5.3}$$

where $\varepsilon_{T_a,i-1}$ is the error associated to the (i-1)-th weather sample and $\varepsilon_{rnd,i}$ is randomly sampled by assuming a normal distribution with zero mean and standard deviation equal to 3.5 K. This approach allows to describe the measurements error in a reasonable way, since it takes into account the precedent meteorological data and the thermal inertia of atmosphere, hence avoiding unphysical temperature steps. The resulting σ_{T_a} value ($\sim$ 2.5 K) can be considered acceptable, according to the previously cited T_a standard deviation estimates (Haiden et al. (2015)).

5.2 Case studies definition

All the optimization approaches investigated in this Chapter were based on the assumption of working with an open pond whose surface area S was 100 m^2. In this Chapter different inlet temperature T_{in} values were tested (20 °C, 25 °C and 30 °C) in order to evaluate their effect on final productivity. The initial conditions used are the same as defined in Chapter 3. The case studies compared in this Chapter are the following ones:

- **Base case**. The base case optimization was based on completely accurate weather data, hence guaranteeing the maximum available productivity at the specific period considered.

- **3σ case**. The 3σ case optimization was based on productivity maximization of function P_{net} (kg) defined in base case through the implementation of a dynamic constraint, as described in the following equation:

$$\begin{array}{ll} \max_{\mathbf{q}} P_{net} = & \max_{\mathbf{q}} \int_0^{\tau} (G(x_b, H_s, T_p) - R(x_b, T_p)) V dt \\ \text{s.t.} & \\ & D_{death} = \int_0^{\tau} (\max(T_p(t) - C_{death}(t), 0)) dt \leq 0, \\ & 0 \leq q_i(t) \leq q_{\max} \\ & l_{p,\min} \leq l_p(t) \leq l_{p,\max} \\ & \dot{\boldsymbol{\xi}} = \mathbf{g}(\boldsymbol{\xi}, \mathbf{q}, t), \quad \boldsymbol{\xi}(0) = \boldsymbol{\xi}_0 \end{array}, \tag{5.4}$$

  where D_{death} is the integral of threshold violation and C_{death} is the implemented constraint on $T_p(t)$, defined as:

$$C_{death} = T_{max} - (T_{p_\sigma}(t) - T_p(t)), \tag{5.5}$$

  where T_{max} is the maximum allowable pond temperature for microalgae growth and $T_{p_\sigma}(t)$ is the dynamic profile obtained for $T_p(t)$ by adding a 3σ error to the extracted

H_s and T_a meteorological values ($3\sigma_{H_s}$ and $3\sigma_{T_a}$, respectively. See section 2.1.4 for errors calculation methodology). In other words, the upper constraint on T_{max} was manipulated in order to get a dynamic constraint that varies depending on the difference between the worst possible value $T_{p_\sigma}(t)$ obtained by adding $3\sigma_{H_s}$ and $3\sigma_{T_a}$ to H_s and T_a extracted values and the predicted pond temperature T_p(t). As an example, Figure 5.1 reports (at $T_{in} = 20$ °C) solar irradiance data modification in July case with respect to base case. The two reported graphs show how the proposed optimization 3σ approach always simulates inaccurate weather solar irradiances that are greater than or equal to base case values. In this way, the 3σ approach takes into account all potential T_{max} threshold excess during cultivation whenever partially cloudy days are predicted but, in fact, sunny days occur.

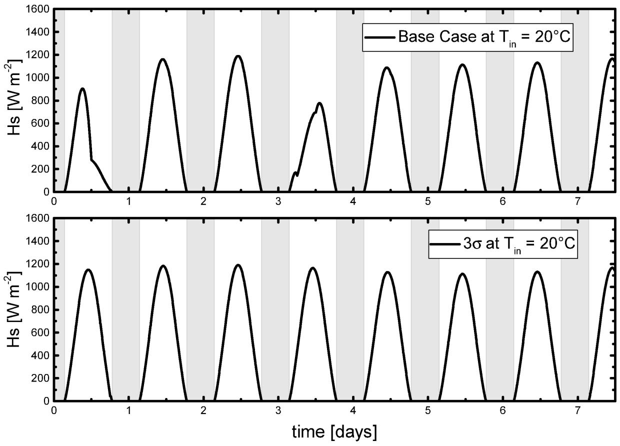
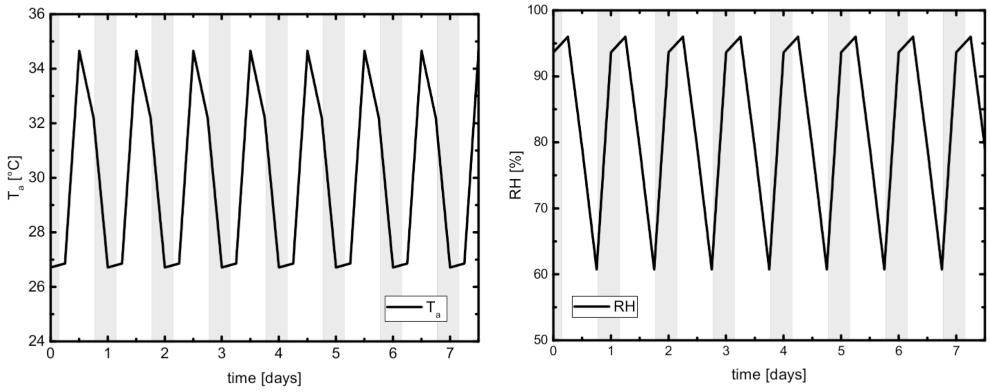


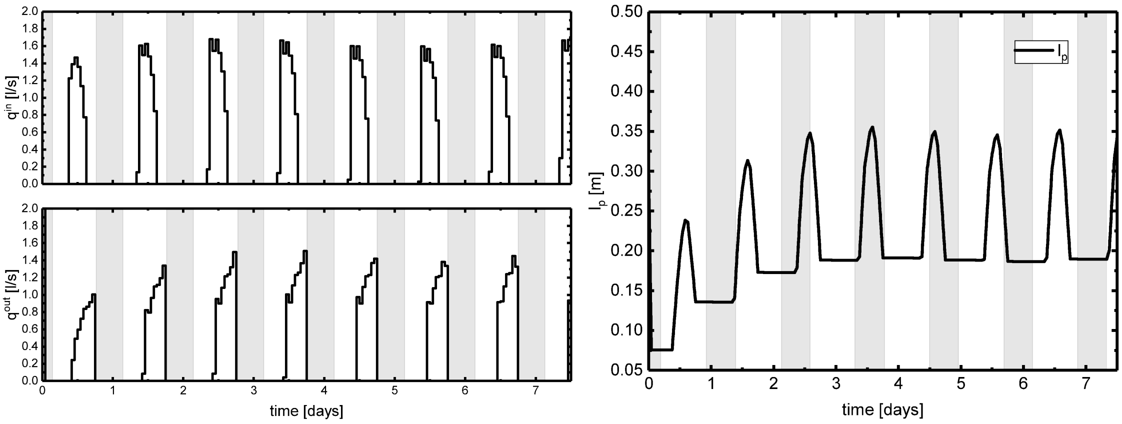
Figure 5.1: *Solar irradiance profile for base case and 3σ case in July at T_{in} = 20 °C. (The background is colored in white at daytime and in grey at nighttime.)*

- **Conservative case**. The conservative case study tackles the problem of weather data uncertainty by assuming that the system has always to cope with the worst growth conditions all along the cultivation time window. The resulting control strategy is iteratively repeated for the whole season, without considering actual meteorological conditions. In other words, the productivity function P_{net} defined in the base case was optimized for the 'worst case' scenario. This 'worst case' meteorological scenario was generated through the daily iteration of the same pattern for both T_a and RH (see Fig. 5.2). In particular, the new T_a and RH profiles were based on the pattern iteration of the day (selected from the first week of July 2012) at which T_p reached its maximum at noon. The wind velocity v_w was kept constant to 0.3 m s^{-1} (minimum value greater than zero in July), whereas both CC and v_r were set equal to zero. The 'worst case' weather data defined above was used to determine the optimal control manipulation (see Fig. 5.3) for all the combinations of T_{in} (20 °C, 25 °C and 30 °C) and all the months (June and July) tested in this work.



(a) (b)

Figure 5.2: *(a) T_a and (b) RH meteorological profiles used for the conservative case. The background is colored in white at daytime and in grey at nighttime.)*



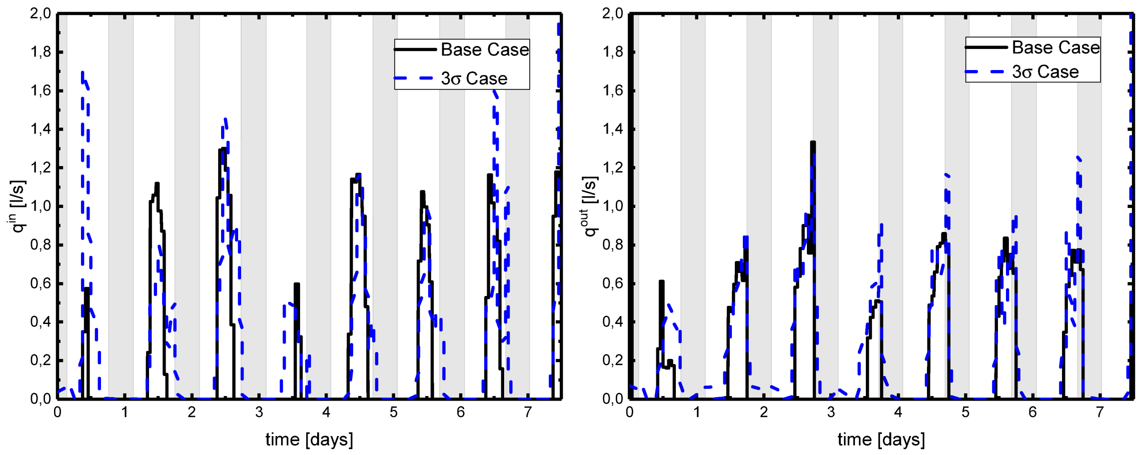
(a) (b)

Figure 5.3: *July: (a) q^{in}, q^{out} and (b) l_p profiles for the conservative case. (The background is colored in white at daytime and in grey at nighttime.)*

5.3 Productivity and water demand comparison between the control approaches

The results presented and discussed in this paragraph are related to the final productivity and water demand associated to the different control approaches described in the previous section. For a complete view of the resulting dynamic profiles of q^{in} and q^{out} obtained for each case study, we refer to Appendix C. In this paragraph the 3σ and the conservative case studies are compared in terms of net productivity P_{net} and water demand WD. All weekly data of P_{net} and WD related to the base case (perfect weather forecasts) are reported in Table 5.6 as reference optimal values, in order to evaluate the drawbacks of each proposed approach with respect to the theoretical optimum. The two control approaches were compared for the months of June and July (2012, Nice). In fact, critical pond temperature values ($T_p \geq T_{max}$) can be reached only during the hottest months of the year in case of inaccurate weather forecasts, hence compromising the most productive period for cultivation (winter and spring weather conditions do not generally present possible risk of microalgae death caused by too high a temperature). Remember that the slight productivity losses reported for 3σ and conservative case are only theoretical biomass losses; in fact, it should be noted that not taking into account weather data inaccuracy leads to cell death, compromising the entire cultivation system.

3σ case. The 3σ case studies were all implemented through the procedure described in the previous paragraph. The resulting constrained optimizations were all based on re-adapting q^{in} and q^{out} values in order to get both maximum productivity and feasible profiles ($T_p \leq T_{max}$) of pond temperature in case of high weather data uncertainty (see, for example, the comparison of q^{in} and q^{out} weekly profiles in July at T_{in} = 20 °C in Fig. 5.4). In fact,



(a) (b)

Figure 5.4: *July, T_{in} = 20 °C: (a) q^{in} and (b) q^{out} profiles for base case and 3σ case. (Plain line: base case; dashed line: 3σ case. The background is colored in white at daytime and in grey at nighttime.)*

the most impactful constraining action was shown at days with the highest differences of solar irradiance with respect to the base case (as an example, the days 1 and 4 of July case represented in Fig. 5.1). This behavior is well represented both for June and July in Figures 5.5 and 5.6), where the dynamic constraint was shown to operate on T_{max} threshold during all the cultivation period at different intensities. The strongest corrective action was reached in correspondence with high forecasts error probabilities (around noon, in correspondence to inaccurately predicted cloudy days).

The resulting weekly productivity values (at T_{in} = 20, 25, 30 °C in June/July) reported in Table 5.6 show that 3σ approach always gave lower results in terms of productivity with respect to base case. Nevertheless, productivity loss was limited to maximum -10% (for June, at T_{in} = 20 °C). The proposed method, in fact, maintained similar q^{in} and q^{out} profiles for sunny days with respect to base case (due to low temperature differences between real and inaccurate predictions, hence low threshold dynamic modifications). Nevertheless, the increase of q^{in} and q^{out} peaks in 3σ (due to high temperature differences between real and inaccurate predictions) during inaccurately predicted cloudy days slightly sacrificed optimal biomass growth conditions in order to guarantee acceptable T_p values through higher culture replacement.

The productivity differences between base case and 3σ case decreased as T_{in} increased, especially in June. This behavior can be explained by considering that the 3σ optimum profiles at high T_{in} values were more similar to the optimum base case profiles than the ones obtained at low T_{in}. At high T_{in} values, in fact, optimal q^{in} assumed higher values around noon, hence increasing the pond depth and, consequently, the thermal inertia of the system. This behavior reduced the difference between T_p and $T_{p,\sigma}$ and, consequently, the dynamic constraint action on the maximum threshold.

The water demand for all 3σ case studies was higher than the one obtained with base cases (up to 158% for T_{in} = 25 °C in June). This behavior was due to the fact that each 3σ case required higher (or equal) q^{in} daily peaks in order to guarantee process feasibility. Furthermore, WD increases with T_{in}; in fact, the higher the inlet temperature, the higher the inflow rate required to cool down the system.

The weekly average concentration of the outlet flowrate was monitored for each T_{in} due to its importance for the downstream processes. High concentration, in fact, reduce the separation costs after the cultivation phase. Table 5.6 shows that the weekly average concentration decreased with T_{in} for both June and July.

Conservative case. Conservative case studies allow to iteratively use the same daily control strategy without considering the necessity of frequent weather forecasts update. Nevertheless the resulting weekly productivity values (at T_{in} = 20, 25, 30 °C in June/July) reported in Table 5.6 show that conservative case approach always gave lower results in terms of productivity with respect to both base case (up to -19% with T_{in} = 20 °C in June) and 3σ case (up to -10% with T_{in} = 20 °C in June). The relative productivity differences between base/3σ case and conservative case decreased as T_{in} increased. The water demand for all conservative case studies was higher than the one obtained with base cases (up to 191% for

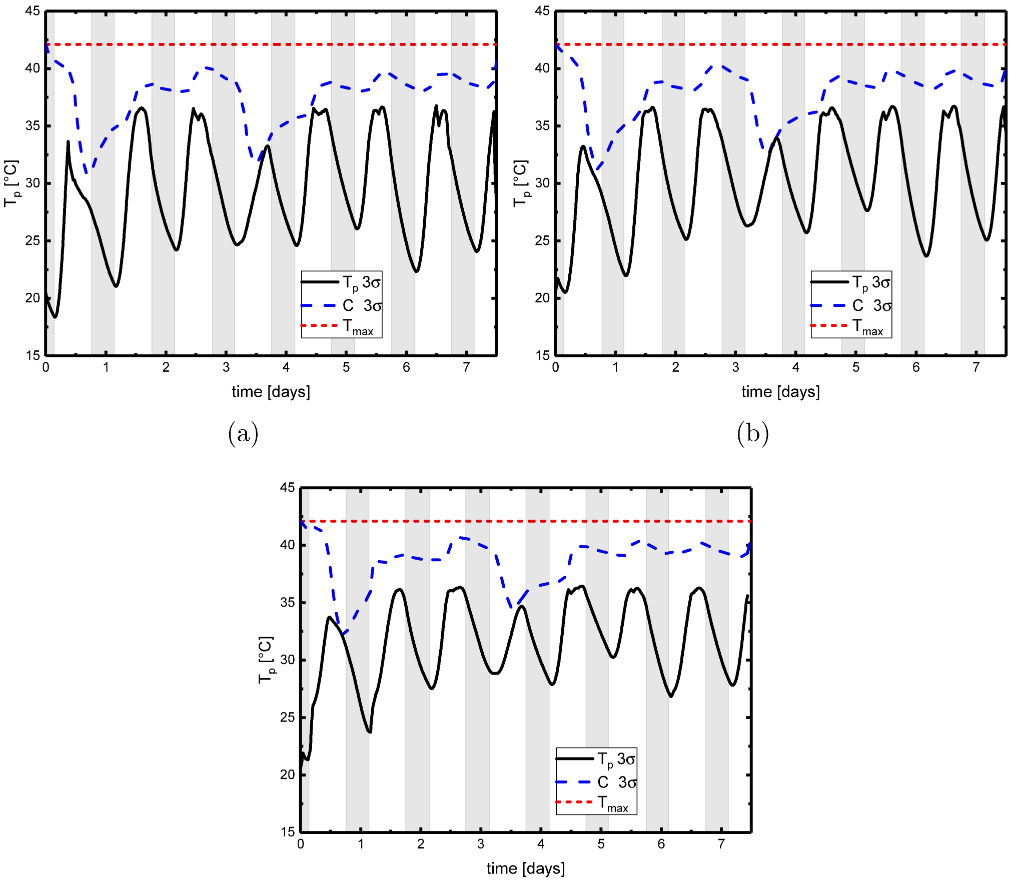
$T_{in} = 30$ °C in June) and 3σ (up to 73% for $T_{in} = 30$ °C in June). Furthermore, WD increased with T_{in} as described for 3σ case studies; in fact, the higher the inlet temperature, the higher the inflow rate required to cool down the system. The weekly average concentration of the outlet flowrate followed the same behavior of the 3σ approach. In fact, Table 5.6 shows that the weekly average concentration decreased with T_{in} for both June and July. Nevertheless, the obtained weekly average concentration values were lower than the values obtained by applying the 3σ approach.

Table 5.5: *Evaluation of the number of scenarios at which $T_p \geq T_{max}$ for base case, 3σ and conservative case at different T_{in} values (June and July, 2012).*

T_{in}	Case studies	$T_p \geq T_{max}$ July	 June
20 °C	base case	62/200	10/200
	3σ	0/200	0/200
	conservative case	0/200	0/200
25 °C	base case	13/200	8/200
	3σ	0/200	0/200
	conservative case	0/200	0/200
30 °C	base case	4/200	9/200
	3σ	0/200	0/200
	conservative case	0/200	0/200

In summary, 3σ case study (with $T_{in} = 20$ °C) gave the best results in terms of productivity loss minimization (with respect to base case), water demand and high biomass concentration required by downstream processes.

The optimization results obtained in this work were compared in terms of operation feasibility. In order to do that, a stochastic simulation of 200 scenarios was conducted for each combination tested (base case, 3σ and conservative case at different T_{in} (20, 25, 30 °C) in June or July). Each scenario represented the model response at different weather data, which was generated by adding a random error (sampled by a normal distribution $N(0,\sigma_{H_s})$ and $N(0,\sigma_{T_a})$) to each H_s and T_a 6-hourly value, respectively. All 200 scenarios related to a specific case study were simulated by implementing the control variables profiles obtained by its optimization (q^{in} and q^{out} dynamics is reported in Appendix C for each case study). In Table 5.5 the scenarios at which feasibility was not reached were reported for each case study. The results show how both the two proposed methodologies (3σ and conservative case) allowed operation feasibility for all the combinations tested. Moreover, Table 5.5 shows that, optimizing the process by assuming ever higher T_{in} values increases process feasibility, even avoiding the implementation of advanced tools to cope with weather inaccuracy (base case). This result can be explained by considering that the higher the inlet temperature, the higher the water amount required by the optimizer. Consequently, both the cultivation volume and the thermal inertia of the system increase, hence guaranteeing more stable temperature conditions in case of wrong predictions.



(a)

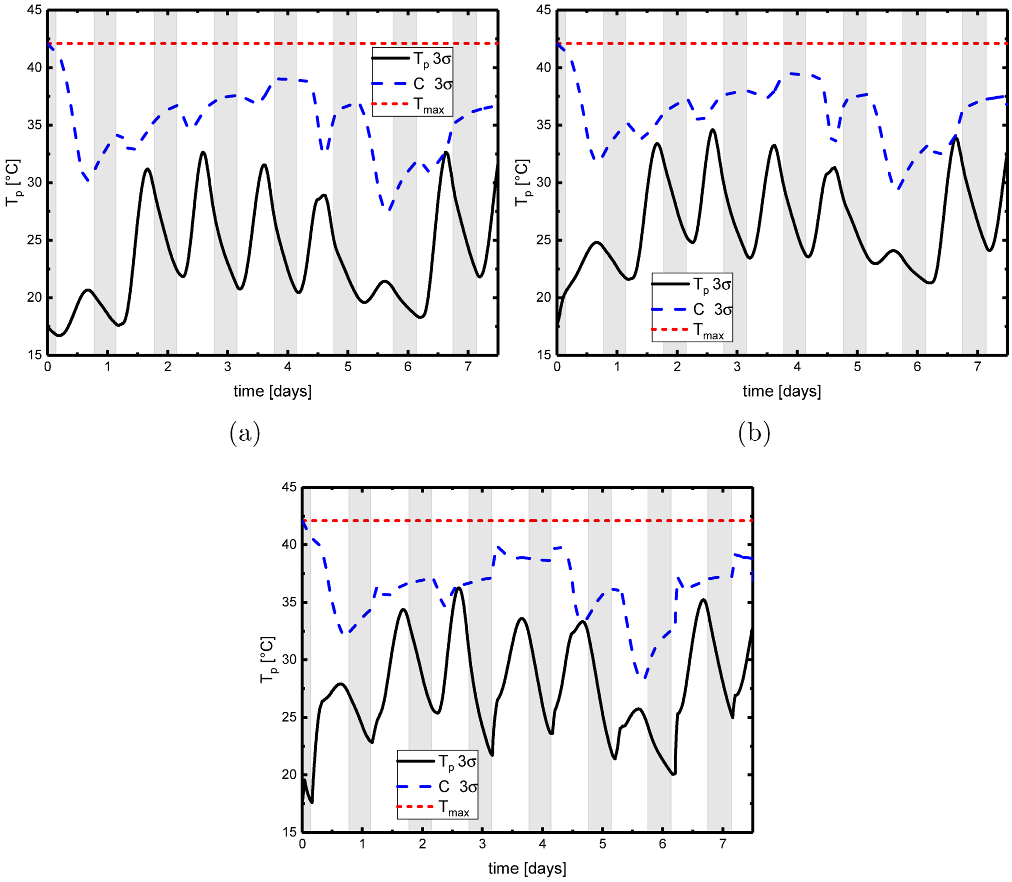
(b)

(c)

Figure 5.5: *July: (a) T_{in} = 20 °C, (b) T_{in} = 25 °C and (c) T_{in} = 30 °C. (Plain line: pond depth; dashed line: dynamic constraint; dot line: T_{max}. The background is colored in white at daytime and in grey at nighttime.)*

Table 5.6: *Productivity, water demand and average concentration of the outlet flowrate for the base case, the 3σ case and the conservative case at different T_{in} values and at different months (June and July, 2012).*

T_{in}	Case studies	Productivity (kg·week^{-1})		Water demand (m^3·week^{-1})		Avg. concentration (kg·m^{-3})	
		June	July	June	July	June	July
20 °C	base case	20.58	25.83	68	123	0.44	0.28
	3σ	18.43	25.36	170	151	0.16	0.20
	conservative case	16.59	24.58	218	218	0.12	0.16
25 °C	base case	20.54	25.82	84	152	0.36	0.23
	3σ	19.44	25.54	217	183	0.13	0.19
	conservative case	18.32	25.07	255	255	0.11	0.13
30 °C	base case	20.70	25.90	126	237	0.24	0.15
	3σ	20.41	25.76	213	270	0.14	0.13
	conservative case	19.61	24.32	368	368	0.08	0.09



(a) (b)

(c)

Figure 5.6: *June: (a) T_{in} = 20 °C, (b) T_{in} = 25 °C and (c) T_{in} = 30 °C. (Plain line: pond depth; dashed line: dynamic constraint; dot line: T_{max}. The background is colored in white at daytime and in grey at nighttime.)*

5.4 Cost analysis

The different control strategies investigated in this work were compared in terms of process cost per biomass produced. The proposed analysis was based on the assumption that the cultivation system was part of the main process schematized in Fig. 5.7. In particular, we assumed identical open ponds working in parallel ($N_{pond} = 50$) through the same optimal control logic. This choice was done in order to simulate an outdoor large-scale plant of 5000 m^2. The proposed flowsheet shows a preliminary set of heat exchangers ($N_{he} = 5$), used to set the inlet medium flowrates at T_{in} for each open pond (the 1:10 ratio between the number of heat exchangers and the number of ponds was chosen to reduce the number of the heat exchanger units required, hence reducing the final cost). The flowrates processed in the heat exchangers consist of the misture recycled from both the settler and the centrifuge and a small amount of fresh water. In this work we assumed that these flowrates enter the heat exchangers at T_p. After the cultivation phase all the culture was harvested in a single settling vessel and then treated in a final centrifuge unit. The installation costs of the open pond was not considered due to its independence by the inlet and outlet flowrates. Cost of nutrients and water was neglected, too. The cost related to flocculant agents needed in the settling unit was neglected due to its low significance with respect to the other cost items. The management and installation costs of the final centrifuge unit was based on the flowrate coming out from the settler. Furthermore, it was assumed to reach a final dry biomass percentage equal to 20% (Ramos Tercero et al. (2014)).

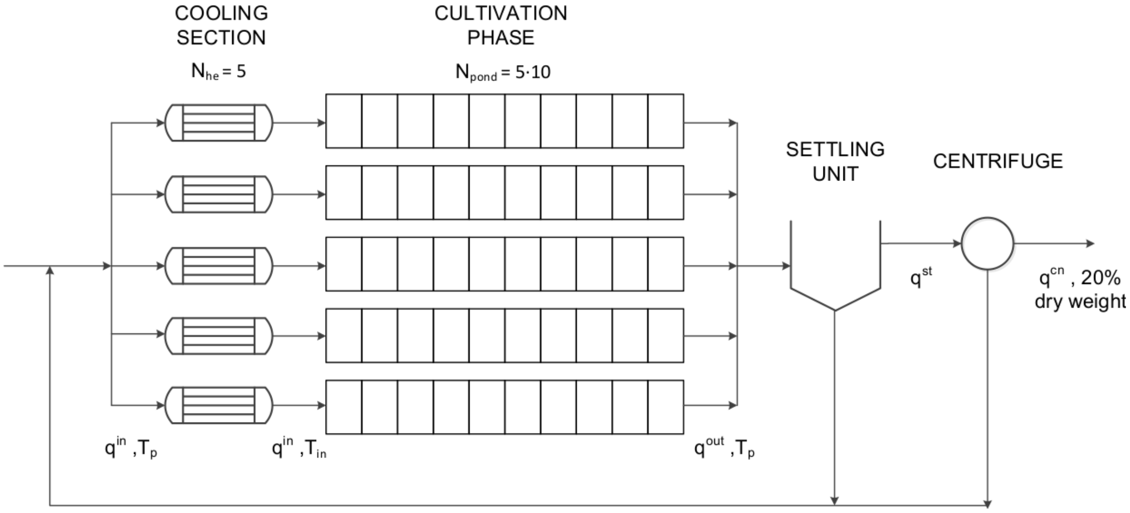


Figure 5.7: *Process flowsheet used for cost analysis*

Since the two control variables to be optimized in this work were the inflow and outflow rates q^{in} and q^{out}, the economic analysis only focused on the process costs that directly depended on these two control variables. The comparison between the costs associated to the different optimization strategies proposed in this work was therefore limited to the process cost component C_v (\$ kg^{-1}), defined as:

$$C_v = \frac{TC_{he} + TC_{st} + TC_{cn}}{N_{pond}(P_{net}/\tau)}, \tag{5.6}$$

where TC_{he} (\$ kg^{-1} week^{-1}) is the total cost associated to the heat exchanger, TC_{st} (\$ kg^{-1} week^{-1}) is the total cost associated to the settler and TC_{cn} (\$ kg^{-1} week^{-1}) is the total cost of centrifuge. The recycle pump cost is not considered here due to its negligible contribution to the total process cost. In other words, C_v represents the difference between the total process cost and all the cost components that are not directly affected by the resulting control strategies (such as the installation cost of the pond, the nutrient cost and the downstream processes not represented in Fig. 5.7). In this paragraph the detailed correlations used to calculate the terms TC_{he}, TC_{st} and TC_{st} in Eq.5.6 are reported. Firstly, the total cost of a generic unit operation TC_{uo} (\$ week^{-1}) was calculated through the following equation:

$$TC_{uo} = MC_{uo} + \frac{IC_{uo}\, CCF}{N_{week}}, \tag{5.7}$$

where MC_{uo} is the manufacturing cost (\$ week^{-1}), IC_{uo} is the installation cost (\$), CCF is the capital change factor (set to 0.333 year^{-1}) and N_{week} is the number of weeks the plant operates per year (set to 39 weeks year^{-1} by excluding winter from the analysis).

5.4.1 The heat exchangers

The heat exchangers manufacturing cost MC_{he} (\$ week^{-1}) was calculated as:

$$MC_{he} = N_{pond} C_{rw} \frac{\int_0^\tau q^{rw}\, dt}{\zeta_\tau \tau}, \tag{5.8}$$

where C_{rw} is the cost of refrigerated water (0.313 \$ m^{-3}, Ramos Tercero et al. (2014)), q^{rw} is the volumetric flow rate of refrigerated water (m^3 s^{-1}) and ζ_τ is the conversion factor used to get results on a weekly basis (1/604800 week s^{-1}). The required amount of q^{rw} was defined by the following equation:

$$q^{rw} = \frac{q^{in}(T_p - T_{in})}{\Delta T_{rw}}, \tag{5.9}$$

where ΔT_{rw} (°C) is the temperature difference of the cold side of the heat exchanger (the cold fluid was assumed to enter the heat exchanger at 5 °C and exit at 15 °C). The heat exchanger installation cost MC_{he} (\$) was expressed from the following correlation (Douglas (1988)):

$$IC_{he} = N_{he} \frac{M\&S}{280} 101.3(2.29 + F_c)(\zeta_A A_{he})^{0.65}, \tag{5.10}$$

where N_{he} is the number of heat exchangers (set to 5), $M\&S$ is the Marshall & Swift cost index (equal to 1576.6, see Douglas (1988)), F_c is a corrective factor (set to 1) and ζ_A is the corrective index used to convert area values to ft^2 ($10.76\ \mathrm{ft}^2\ \mathrm{m}^{-2}$). The heat exchanger area A_{he} (m^2) was expressed as:

$$A_{he} = \frac{N_{pond}}{N_{he}} \frac{Q_{he}}{U_{he} \Delta T_{ml}}, \tag{5.11}$$

where U_{he} is the global heat transfer coefficient ($625\ \mathrm{W\ m^{-2}\ K^{-1}}$) and ΔT_{ml} (°C) is the logarithmic mean temperature difference, calculated as:

$$\Delta T_{ml} = \frac{(T_{opt} - T_{in}) - \Delta T_{rw}}{\ln[(T_{opt} - T_{in})/\Delta T_{rw}]}. \tag{5.12}$$

The term Q_{he} in Equation 5.11 represents the heat exchanger duty (W), obtained through the following correlation:

$$Q_{he} = \rho_w c_{p_w} q^{he} (T_{opt} - T_{in}), \tag{5.13}$$

where q^{he} represents a volumetric flow rate ($\mathrm{m^3\ s^{-1}}$) defined as:

$$q^{he} = \min(q^{in,max}, q^{buf,max}). \tag{5.14}$$

The variable $q^{in,max}$ in Equation 5.14 is the maximum value of q^{in} during the cultivation period whereas $q^{buf,max}$ is defined as:

$$q^{buf,max} = \max_{1 \le i \le n_{day,w}} \left(1.3 \frac{\int_{t_{1,i}}^{t_{2,i}} q^{in}\, dt}{\zeta_{\tau}(t_{2,i} - t_{1,i})} \right), \tag{5.15}$$

where $t_{1,i}$ is the first time at which q^{in} assumes non zero values during the i-th day, $t_{2,i}$ is the last instant at which q^{in} assumes non zero values during the i-th day and $n_{day,w}$ is the number of days used in the simulation/optimization. The coefficient 1.3 in Eq.5.15 was used to increase the value of the involved flow rates for conservative heat exchanger design.

5.4.2 The settling unit

The settler was used to harvest the culture from each open pond (in this work we assumed to have 50 open ponds that operate in the same way). The outlet biomass concentration was assumed to be five times higher than the inlet culture medium concentration. The decanter area A_{st} (m^2) was calculated as:

$$A_{st} = \frac{V_{max} N_{pond}}{h_{st}}, \tag{5.16}$$

where V_{max} (m^3) is the maximum daily volume of culture extracted by the open pond during the cultivation period and h_{st} represents the decanter depth (fixed at 6 m). The obtained value of A_{st} was used to estimate the installation cost of the settler IC_{st} ($) through the following empirical correlation (Sharma et al. (2013)):

$$IC_{st} = \alpha_{1,st} (\zeta_A A_{st})^2 + \alpha_{2,st} \zeta_A A_{st} + \alpha_{3,st}, \tag{5.17}$$

where $\alpha_{1,st}$ (-0.0005 \$ ft^{-2}), $\alpha_{2,st}$ (86.89 \$ ft^{-1}) and $\alpha_{3,st}$ (182801 \$) are empirical coefficients.

5.4.3 The centrifuge

Since the biomass concentration of the settler outlet flow q^{st} (m^3 s^{-1}) was not high enough for the downstream processes, one centrifuge unit was implemented to reach the 20% of dry biomass (Molina Grima et al. (2003)). The settler outlet flow q^{st} was defined as:

$$q^{st} = \frac{V_{max} N_{pond}}{\gamma_x t_{op}}, \tag{5.18}$$

where γ_x is a concentration index used to scale the centrifuge volume with respect to the settler unit (set to 5) and t_{op} is the assumed operation duration (57600 s). The installation cost of the centrifuge IC_{cn} was calculated as:

$$IC_{cn} = \beta_{1,cn}(\zeta_q q^{st})^3 + \beta_{2,cn}(\zeta_q q^{st})^2 + \beta_{3,cn}(\zeta_q q^{st}) + \beta_{4,cn}, \tag{5.19}$$

where ζ_q is the flowrate conversion factor (15850.3 gal s min^{-1} m^{-3}). $\beta_{1,cn}$ (0.0133 \$ min^3 gal^{-3}), $\beta_{2,cn}$ (12.685 \$ min^2 gal^{-2}), $\beta_{3,cn}$ (5635.3 \$ min gal^{-1}) and $\beta_{4,cn}$ (411407 \$) are empirical coefficients. Finally, the manufacturing cost MC_{cn} was calculated as:

$$MC_{cn} = \frac{19829(\zeta_q q^{st})^{0.4168}}{N_{week}}. \tag{5.20}$$

5.4.4 Economic comparison between the control approaches

In this paragraph the 3σ and the conservative case studies were compared through the evaluation of the cost index C_v previously defined. The two control approaches were tested at different inlet temperatures (20 °C, 25 °C and 30 °C) for the two sunniest, most productive and water demanding months (June and July 2012, Nice) considered in this work. The results are reported in Table 5.7. In particular, q^{he} was calculated as the minimum value between the maximum weekly q^{in} peak and the function $q^{buf,max}$ described in Eq.5.15.

July. Table 5.7 shows that the lowest C_v value obtained with the 3σ approach was the one at T_{in} = 20 °C (C_v = 17.01 \$ kg^{-1}), whereas the highest was C_v = 18.23 \$ kg^{-1} at T_{in} = 30 °C (+ 7%). The conservative case does not show high cost differences between the simulations at T_{in} = 20 °C and T_{in} = 25 °C, but both these two conservative case studies gave higher cost estimations (up to 25% with respect to 3σ approach). Conversely, the simulation at T_{in}= 30 °C for the conservative case presented the highest cost increase (+57% if compared to the C_v value obtained in the 3σ case at T_{in} = 30 °C).

June. Table 5.7 shows that all C_v values obtained in the conservative case were higher than those resulting from the use of 3σ approach. The highest C_v values were obtained at T_{in} = 30 °C, similarly to July results.

In summary, the two proposed approaches allow to guarantee both cultivation process feasibility (in case of inaccurate forecasts) and high productivity. In fact, the optimal productivity values for both the two approaches are only slightly lower than the ones achievable under the assumption of perfect meteorological data knowledge, especially if 3σ approach is used. Unfortunately, both the proposed methods are more costly and water consuming with respect to the base case; nonetheless, the 3σ approach can be considered the best option both in terms of costs and water demand with respect to the conservative case for the two months considered. This results can be justified by the fact that 3σ approach is based on the dynamic adaptation of the temperature threshold, locally adjusted by following the inaccuracy estimates generated by varying weather data. Conversely, the conservative approach is more water, energy and cost demanding, due to the recursive control logic resulting from the assumption of the daily iteration of the worst weather conditions for the whole cultivation period. Nevertheless, the conservative approach is the best solution if the user is looking for the simplest control management guaranteeing process feasibility. Finally, the results obtained for all the two approaches show that, the lower the temperature of the inlet flowrate, the lower the water demand required by the optimization and, consequently, the associated process cost. This result is based uniquely on the temperature values tested in this paper (20 °C, 25 °C and 30 °C); the extrapolation of this statement to lower T_{in}s should therefore be assessed through specific simulations.

Table 5.7: *Evaluation of C_v (\$ kg^{-1}) for the 3σ and conservative case at different T_{in} values (June and July, 2012).*

T_{in}	Case studies	July	June
20 °C	3σ	17.01	21.08
	conservative case	20.78	29.99
25 °C	3σ	17.13	20.84
	conservative case	21.08	27.68
30 °C	3σ	18.23	19.61
	conservative case	28.50	30.89

5.5 Conclusions

Mathematical models able to describe microalgae growth in open ponds through meteorological data are valuable tools for increasing cultivation phase productivity and giving hints for practical operation. Nevertheless, the results obtained through this kind of models could be misleading since they are based on the assumption of perfect weather data. Firstly, two advanced optimal methods were proposed in this work in order to cope with uncertain weather forecasts, leading to potential critical conditions for microalgae growth. Both the two approaches guaranteed cultivation feasibility; moreover, the discrimination between the two approaches led to the selection of the best approach (in particular, 3σ case) in terms of productivity loss minimization, water demand reduction, optimal conditions for downstream processes and total process costs. Secondly, the two advanced optimal methods were tested at different inlet flowrate temperature conditions; T_{in} variable has, in fact, a key role on control magnitude peaks and, consequently, on water demand and productivity. The final results showed that the best options was to operate at the lowest T_{in} temperature tested (20 °C). Further research needs to address the optimal control conditions and T_{in} temperature required to maximize a new objective function embedding both productivity and total process costs.

Chapter 6

Conclusions and future perspectives

Microalgae are one of the most promising renewable feedstocks to cope with fossil sources depletion and provide both high-value innovative compounds and biofuels. However, the current assessment on practical feasibility and economics/environmental benefits is generally limited by many uncertainties, especially regarding the biomass productivity that can actually be reached at full-scale. In this context, mathematical modelling can be of great help for developing our current knowledge on microalgae growth dynamics and optimizing design, operation and control of microalgae culture systems. This Thesis work aimed at exploring modelling issues with both a microscopic and a macroscopic perspective. From a microscale perspective, we proposed an extension of an existing state-model in order to deepen our current knowledge about the whole electron transport process involved in photosynthesis. From a macroscale point of view, we considered an 'in silico' representation of outdoor systems by coupling growth and temperature models to meteorological forecasts. We precise that these two research areas were kept disconnected in this work, though they may be merged in a unique multiscale model. The main achievements obtained with these two different approaches and some hints for future research will be discussed in the following. A final investigation on new MBDoE procedures based on online redesign of experiments (see Appendix A) has been reported in this Thesis work, though no direct link to microalgae modelling is considered. Nevertheless, the current benefits of standard MBDoE techniques for microalgae model calibration/validation (see Bernardi et al. (2016) work) and the possibility to use the proposed modified version to reduce otherwise long microalgae cultivation experiments have a great potential in improving future model-based design of experiment.

6.1 Microscale approach

The first objective of this Thesis was to propose an extension of an existing semi-mechanistic state-model representing three distinct processes acting on PSII at different time scales (photoproduction, photoinhibition and Non Photochemical Quenching (NPQ)) and at different light intensities through fluorescence fluxes mathematical description. The objective was to use this model as a starting point for simulating the entire electron transport process triggered by PSII in order to have a complete view of the most significant photosynthetic mechanisms. In the following the main achievements are summarized:

- The proposed model proved to be a good candidate for describing PSI redox dynamics. The model was calibrated through fast spectroscopic analysises performed in vivo using a Joliot-type spectrophotometer. All the model parameters were estimated in a statistically satisfactory way, hence guaranteeing local model identification; moreover, the model is able to fit the experimental data in an accurate way;
- The model was validated against LEF and CEF experimental data in order to discriminate between two different models that differed in the representation of linear electron flow. The resulting best model is the one representing LEF reduction by implementing lumen acidification through PQ transport rate regulation;
- Simulations of the dynamics of all the protein complexes involved in the electron has been made with the proposed validated model. However, the final state of some protein complexes does not correspond to its initial redox state. These model predictions could not be verified with available experimental data; however, to our knowledge, no literature model has been tested to demonstrate the existence of unkwnown dark phase-driven hysteresis phenomena for over-reduced electron carriers.

The aim of future research should be to propose further investigation on the behavior of all the electron carriers between PSII and PSI through dedicated protocols. A preliminary test could be done by designing a a series of rapid dark/light cycles to test if protein complexes redox state at the beginning of each light phase assume or not different initial values, depending on the light path followed. Then, further work should be done for dedicated absorbance experiment of the electron carriers between PSII and PSI, in order to improve the process description given by the current model and to confirm (or falsify) some of the model predictions that could not be verified with available experimental data. The model has been validated for *Nannochloropsis gaditana* strain; further work is required to assess the model reliability to describe different micralgae strains. New experimental campaigns should be designed in order to collect new data and calibrate/validate the proposed model for each new strain considered. Finally, model structure modifications should be taken into account since the complexes involved in the electron transports vary depending on the microalgae type.

6.2 Macroscale approaches

The second objective of this Thesis was to propose an optimization approach to maximise microalgae productivity in open-pond systems at different weather conditions. The model was built by coupling existing growth/temperature sub-models with real meteorological data. This dynamic model was used to test a new approach for raceways management and underline the benefits of model building activities on practical process optimization and control. Microalgal productivity maximization was achieved by continuously adjusting the open pond depth through both the injection of fresh medium or extraction of culture. In the following the main achievements are summarized:

- The control strategy proposed in this Thesis significantly increased microalgae weekly productivity compared to standard operation at constant depth and dilution rate. The optimal meteorological-based control strategy allowed to increase productivity values up to a factor 2.2 for hot and sunny climates; however, this control strategies induces high dilution rates, hence high water demand, especially in summer;

- The control strategy is mainly based on adapting the raceway pond volume and flowrates to get as close as possible to the optimal temperature for the cultivated strain and prevent cell death due to too high temperatures. The analysis of the rational behind the optimal control revealed that process optimization are always based on two main mechanisms: periodically 'flushing' the culture and controlling the raceway pond depth. The compensation condition is then tracked once temperature is close to the optimal value. Furthermore, we demonstrated that the complex optimal strategy can be reduced to simple guidelines for future practical operation;

- Since the proposed control strategy is highly dependent on weather forecasts, an accurate prediction of meteorological conditions must be available at least for one day to guarantee an efficient process optimization. Daily re-optimization in fact allows to maintain productivity values close to the ideal case based on weekly perfect forecasts, due to the fact that daily control strategy is little influenced by weather conditions related to other days. Inaccurate predictions can be compensated by a more frequent optimization update;

- Uncertain weather forecasts, leading to potential critical conditions for microalgae growth after optimization, can also be compensated by using different techniques to guarantee process feasibility. In particular, the implementation of an additional dynamic constraint on the pond temperature has been proven to be the best option to maintain feasible growth conditions, thus avoiding cell death. This constraint, based on local weather data inaccuracy estimates, allowed to maintain high productivities, with a low increase of water demand and operating costs.

The aim of future research may be to propose the implementation of additional phenomena in the current model; in fact, some important biological mechanisms, such as photoinhibition and photoacclimation, are not considered though their impact on final microalgae productivity may be significant. Furthermore, all the results presented are based on complex optimization runs whose final results could be sub-optimal, given the high degree of freedom of the problem. A multiple shooting approach or advanced (global) optimization techniques may therefore be used in order to check the actual optimization reliability. Finally, since no model mismatch was assumed, future experiments should be performed to validate the model against real/historical data and to estimate the amount of uncertainty generated by each sub-model.

As a perspective, the same optimization strategy should be applied, but by constraining the amount of water that can be used in the process (based for example on the availability on rainwater at the location considered), and assuming that an important fraction of the water can be recycled. This target could also be reached by implementing a multi-objective optimization in order to investigate the control strategy required to get the best trad-off between productivity maximization and water demand minimization.

The preliminary process cost evaluation made in Chapter 5 to discriminate between different feasibility strategies showed that operating with inlet flowrates at low temperatures is recommended to minimize the water demand. In fact, the lower the inflow temperature T_{in}, the lower the amount of water needed to cool down the system to avoid critical temperatures in the pond. However, cooling down water requires a refrigeration system, hence increasing the process cost. Further research could focus on defining a more detailed process flowsheet and evaluate the costs associated to each process unit. An interesting research direction, in this case, could be to implement manufacturing and operation costs in the optimization framework and adding T_{in} as control variable, by maintaining acceptable computational time for future optimization tasks.

Finally, a third challenge could be to implement the optimal control strategy to an actual MPC framework, in order to assess the reliability of the 'in silico' results obtained in this Thesis to real open pond systems. Furthermore, advanced model predictive control techniques, such as Adaptive MPC, should be considered in order to cope with the ideal assumption of 'no model mismatch'.

Appendix A

Information driven approach for online model-based redesign of experiments

In this Chapter[1] a new online model-based design of experiments technique is discussed and tested for two different case studies. The main objective is to exploit the progressive increase of information resulting from running experiments in case of high parametric mismatch. The proposed methodology allows determining when to redesign the experiment in an automatic way, thus guaranteeing that an acceptable increase in the information content has been achieved before proceeding with the intermediate estimation of the parameters and the subsequent redesign of the experiment. Although the effectiveness of the proposed experiment design technique is demonstrated through two simulated case studies not related to the Thesis topic, the methodology was presented due to its applicability to microalgae growth models (see, as an example, Bernardi et al. (2016) work).

A.1 Introduction

Model-based design of experiments (MBDoE) techniques (Pukelsheim (1993)) represent a valuable tool for the rapid assessment and development of mathematical models at different levels of the model building procedure (Asprey and Macchietto (2002)), allowing for the maximization of the experimental information in order to reduce time and costs of the model identification task. The effectiveness of the conventional iterative MBDoE procedure (Franceschini and Macchietto (2008)) has been proved in a large variety of applications (Prasad and Vlachos (2008); Galvanin et al. (2009b); Chakrabarty et al. (2013)) but it is greatly limited by the fact that the design activity is affected by the initial available estimates of parameter values; therefore high uncertainty on these values can severely

---

[1] Part of this work is reported in the article by De-Luca R., F. Galvanin, F. Bezzo. (2016). A methodology for direct exploitation of available information in the online model-based redesign of experiments. *Comput. Chem Eng.*, 91:195-205

affect the efficiency of the experimental design (Körkel et al. (2004)). In fact, several approaches were proposed in order to overcome this problem. For example, Houska et al. (2015) proposed an iterative optimal design method that consists of a modified A-criterion weighting the terms information matrix trace calculated through an ad-hoc algorithm. As stated by Mehra (1974), an efficient design procedure may be based on the exploitation of the information as soon as it is generated by the running experiment through an online adaptive input design (AID) strategy. AID strategies have been proposed and applied to linear stochastic control system (Lindqvist and Hjalmarssonn (2001); Gerencsér et al. (2009), where the design task is solved over a preset time horizon after which a new measurement is taken and, consequently, a new estimation is acquired. Extension of this procedure to nonlinear dynamic system is quite a recent achievement (Stigter et al. (2006)) and an online model-based redesign of experiments (OMBRE) approach has been recently proposed as a natural expansion of AID strategies (Galvanin et al. (2008, 2009a)). In OMBRE the manipulated variables dynamic profiles and the sampling points allocation are updated by performing one or more intermediate experiment designs (redesigns). The efficiency of this technique has been tested for a wide range of cases by different research groups both for model validation (Barz et al. (2013)) and model discrimination issues (Schenkendorf and Mangold (2013)), providing a very efficient usage of measurement data and a great improvement with respect to conventional MBDoE techniques. The OMBRE approach has been also recently extended to systems where disturbances and systematic errors may be present by using model updating policies including disturbance estimation procedures which are embedded within the OMBRE strategy (Galvanin et al. (2012)). Unfortunately, OMBRE technique is affected by some limitations, too: firstly, the redesign policy is decided a priori by the user, without any rational criterion related to the achievable information; secondly, OMBRE approach is still affected by the initial parameter uncertainty, especially until a first redesign is done. In this Chapter a novel information-driven redesign optimization (IDRO) is presented, where a robust design approach is applied to the online redesign procedure with the purpose of determining when to redesign the experiment in an automatic and robust way. IDRO is based on a new design concept, based on the maximization of a target profile of the dynamic information profile, which guarantees a reliable increase in the information content before proceeding with the intermediate estimation of the parameters and the subsequent redesign of the experiment. Furthermore, the technique is much less affected by the negative effects of parametric uncertainty, with great benefit in terms of robustness of the whole redesign procedure. The applicability to nonlinear dynamic systems is demonstrated through two simulated case studies: the first one is related to the identification of a fermentation bioreactor model, the second one is based on the identification of a physiological model which describes glucose homeostasis.

A.2 Optimal design of experiments: methodology

A conventional MBDoE procedure (Asprey and Macchietto (2002); Körkel et al. (2004)) aims at decreasing the model parameter uncertainty region by acting on the n_φ-dimensional experiment design vector $\boldsymbol{\varphi}$ and solving the following nonlinear optimization problem:

$$\begin{aligned} & \boldsymbol{\varphi}^{opt} = \underset{\boldsymbol{\varphi}}{\operatorname{argmin}} \left\{ \psi \left[\mathbf{V}_\theta \left(\boldsymbol{\theta}, \boldsymbol{\varphi} \right) \right] \right\} = \underset{\boldsymbol{\varphi}}{\operatorname{argmin}} \left\{ \psi \left[{\mathbf{H}_\theta}^{-1} \left(\boldsymbol{\theta}, \boldsymbol{\varphi} \right) \right] \right\} \\ \text{s.t.} & \\ & f\left(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \mathbf{w}, \boldsymbol{\theta}, t \right) = 0, \quad \hat{\mathbf{y}} = h\left(\mathbf{x}(t) \right) \\ & \mathbf{C} = \mathbf{x}(t) - \mathbf{G}(t) \leq 0 \\ & \boldsymbol{\varphi}^l \leq \boldsymbol{\varphi} \leq \boldsymbol{\varphi}^u \end{aligned} \quad , \tag{A.1}$$

with the set of initial conditions $\mathbf{x}(0) = \mathbf{x}^0$. In Eq. A.1 $\mathbf{V}_\theta$ is the variance-covariance matrix of model parameters and $\mathbf{H}_\theta$ corresponds to its approximated inverse, the dynamic Fisher information matrix; $\mathbf{x} \in \Re^{N_x}$ is the time-dependent state variables vector, $\mathbf{u}(t) \in \Re^{N_u}$ and $\mathbf{w} \in \Re^{N_w}$ are, respectively, the time-dependent and time-invariant control variables (manipulated inputs), $\boldsymbol{\theta} \in \Re^{N_\theta}$ is the model parameters set, and t is time.

The symbol ($\hat{\ }$) is used to specify the estimated value of a variable (or a set of variables): for example, $\mathbf{y}$ represents the vector of measured values of the outputs, while $\hat{\mathbf{y}}$ is the vector of the corresponding values estimated by the model. In Eq. A.1 $\mathbf{C}$ is the N_C-dimensional set of constraint functions expressed through the set $\mathbf{G}(t) \in \Re^{N_C}$ of active constraints on state variables. Equation A.1 represents the n_φ-dimensional set of constraints on design variables, usually expressed by lower (superscript l) and upper (superscript u) bounds on each components of the experiment design vector $\boldsymbol{\varphi}$, constraining the design to a hyperrectangular subspace of the overall design space $\Re^{n_\varphi}$. The experiment design vector $\boldsymbol{\varphi}$ is defined as:

$$\boldsymbol{\varphi} = \left[\mathbf{y}^0, \mathbf{u}(t), \mathbf{w}, \mathbf{t}^{sp}, \tau \right] \tag{A.2}$$

It consists of the N_y-dimensional set of initial conditions for the measured variables $\mathbf{y}^0$, the manipulated input variables $\mathbf{u}(t)$ and $\mathbf{w}$, the total duration of the experiment τ and the N_{sp}-dimensional set of output variables sampling times $\mathbf{t}^{sp}$.

The function ψ in Eq. A.1 is an assigned metric of the variance-covariance matrix of model parameters $\mathbf{V}_\theta$ and represents the design criterion adopted to maximize the expected information content of the experiment as predicted by the model; the most common design criteria are the alphabetical ones, i.e. A-, D-, E-optimal criteria which focus on trace, determinant and maximum eigenvalue of $\mathbf{V}_\theta$ minimization, respectively (Pukelsheim (1993)). The dynamic information matrix $\mathbf{H}_\theta$ for a single experiment is usually expressed by a discrete dynamic form of the Fisher information matrix (Bard (1974)). According to the notation proposed by Zullo (1991), $\mathbf{H}_\theta$ is here defined as:

$$\begin{aligned}\mathbf{H}_\theta(\boldsymbol{\theta},\boldsymbol{\varphi}) &= \mathbf{H}_\theta^0+\sum_{k=1}^{N_{sp}}\sum_{i=1}^{N_y}\sum_{j=1}^{N_y} s_{ij}\left[\frac{\partial\hat{y}_i(t_k)}{\partial\theta_l}\frac{\partial\hat{y}_j(t_k)}{\partial\theta_m}\right]_{l,m=1,\dots,N_\theta} \\ &= \mathbf{H}_\theta^0+\sum_{k=1}^{N_{sp}}\mathbf{M}_k\end{aligned}, \tag{A.3}$$

where s_{ij} is the ij-th element of the inverse of the $N_y \times N_y$ measurement error covariance matrix $\boldsymbol{\Sigma}_y$, $\mathbf{M}_k$ represents the amount of information that can be obtained at the k-th sample and $\mathbf{H}_\theta^0$ is the preliminary information matrix, based on the preliminary statistics about the parametric system at the beginning of the experimental campaign. The conventional MBDoE (Franceschini and Macchietto (2008)) can be described as an iterative loop procedure generally based on the following steps:

1. get prior knowledge on the parametric set value (and its related uncertainty $\boldsymbol{\Sigma}_\theta$);
2. choose an optimal design criterion and design the experiment;
3. start the experiment;
4. at the end of the experimental run, estimate the model parameters;
5. if the desired estimation quality is not reached by the end ofthe experiment, design a new experiment, based on the model parameters estimated in step 4.

A.3 Online model-based redesign of experiments: OMBRE

When OMBRE procedure is used, intermediate parameter estimations are carried out at specific updating times while the experiment is running, in order to exploit the information obtained and use it to partially design the remaining part of the test. The experimental run is thus split into sub-experiments, within which the experiment decision variables are distributed according to the 'a priori' chosen redesign strategy (Galvanin et al. (2009a)). Following this approach, the global design vector of the experiment $\boldsymbol{\varphi}$ can therefore be rewritten as:

$$\boldsymbol{\varphi} = \left[\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_j, \dots, \boldsymbol{\varphi}_{n^{up}+1}\right], \tag{A.4}$$

where n^{up} is the number of updating times, and $\boldsymbol{\varphi}_j$ is the design vector before the j-th update; each component $\boldsymbol{\varphi}_j$ of $\boldsymbol{\varphi}$ could have a different dimension in terms of number of discretized control variables and/or sampling points. Obviously, $\boldsymbol{\varphi}_1$ will be the only component including the initial values of measured variables. For each redesign activity an optimization problem acting on the j-th component of $\boldsymbol{\varphi}$ is conducted in the corresponding time frame; in particular, the optimality condition for $\boldsymbol{\varphi}_j$ is given by:

$$\boldsymbol{\varphi}_j^{opt} = \underset{\boldsymbol{\varphi}_j}{\operatorname{argmin}}\left\{\psi\left[\mathbf{V}_{\theta|j}(\boldsymbol{\theta},\boldsymbol{\varphi}_j)\right]\right\} = \underset{\boldsymbol{\varphi}_j}{\operatorname{argmin}}\left\{\psi\left[\mathbf{H}_{\theta|j}(\boldsymbol{\theta},\boldsymbol{\varphi}_j)\right]\right\} \quad , \tag{A.5}$$

where the information acquired by the j-th redesign can be expressed as a partial contribute to the total dynamic information matrix:

$$\mathbf{H}_{\theta|j}\left(\boldsymbol{\theta},\boldsymbol{\varphi}_j\right)=\tilde{\mathbf{H}}_{\theta}\left(\boldsymbol{\theta},\boldsymbol{\varphi}_j\right)+\sum_{k=0}^{j-1}\tilde{\mathbf{H}}_{\theta|k}\left(\boldsymbol{\theta},\boldsymbol{\varphi}_k\right)+\boldsymbol{\Sigma}_{\theta}{}^{-1}=\tilde{\mathbf{H}}_{\theta}\left(\boldsymbol{\theta},\boldsymbol{\varphi}_j\right)+\mathbf{L} \quad . \tag{A.6}$$

The constant term $\mathbf{L}$ consists on the sum between the preliminary information matrix $\boldsymbol{\Sigma}_{\theta}{}^{-1}$ and the information acquired before the j-th redesign. The symbol $(\tilde{\ })$ indicates that the information matrix refers to a single updating interval, and $\tilde{\mathbf{H}}_{\theta|0}$ is the null matrix. At each given updating time the information is obtained by executing online a parameter estimation session followed by a redesign of the remaining part of the experiment. The detailed OMBRE procedure is therefore based on the following steps (see Fig. A.1):

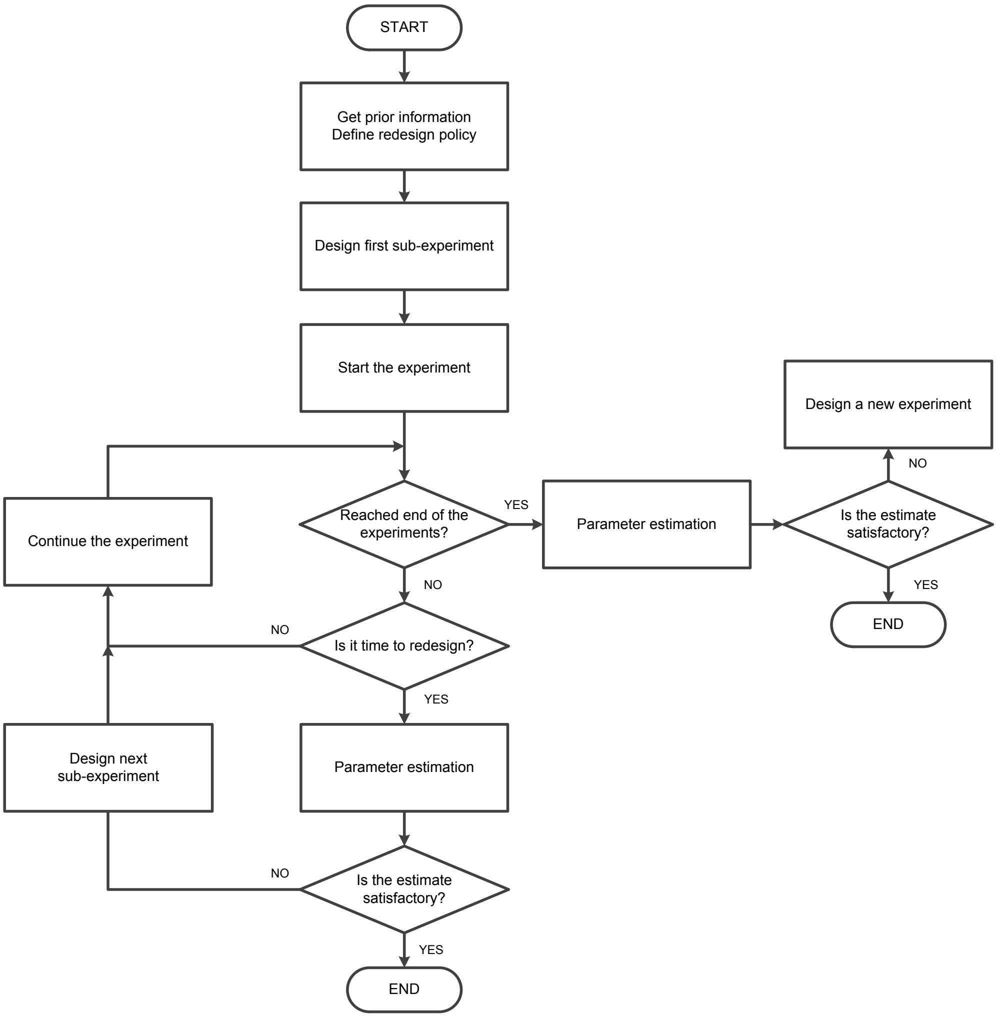


Figure A.1: *Flowchart of OMBRE procedure.*

1. get prior knowledge on the parametric set value (and its related uncertainty $\boldsymbol{\Sigma}_{\theta}$);
2. choose an updating strategy and design the first sub-experiment by calculating A.5;
3. start the experiment;

4. if an updating time is reached, estimate the model parameters:
5. if a statistically sound parameter estimation is achieved, then stop the redesign procedure (and possibly the experiment itself); otherwise:
    - (a) redesign the remaining part of the experiment by using Eq. A.5 with the update on dynamic information given by Eq. A.6; implement the design in the running experiment, and go to step 4.
    - (b) if the desired estimation quality is not reached by the end of the experiment, design a new experiment, based on the model parameters estimated in step 4.

Note that a possible parametric mismatch can be managed by OMBRE only by adjusting the model parameters according to the available observations. The optimality condition and the feasibility conditions will be evaluated within each time frame according to the current value of model parameters.

A.4 Information-driven redesign optimization (IDRO) for model-based design of experiments

In OMBRE methodology the updating strategy is chosen a priori by the user without taking into account any rational criterion related to the achievable information; in order to overcome this issue a new updating strategy is here proposed including: (*i*) the optimization of the information obtainable from the experiment based on a target information and (*ii*) the identification of the best updating time to redesign. The design task, especially at the beginning of the procedure, may be affected by an initial parametric mismatch.

Consequently, in order to preserve the quality of the experiment by poor starting values of the parameters, a robust optimal design approach (Asprey and Macchietto (2002)) based on the expected value of a measurement function of the predicted information is implemented. IDRO methodology is essentially based on a robust A-optimal criterion applied for the worst model parameter, by exploiting one of the most significant statistical indices used to evaluate the parameter estimation precision: the t-value. This index is strictly linked to the information that can be gained during the experimental run, as can be observed from the following equation:

$$t_i = \frac{\theta_i}{t\left(1/2 + (1-\alpha/2), N_{sp} - N_\theta\right) \cdot \sqrt{v_{ii}}}, \quad i = 1, ..., N_\theta \quad , \tag{A.7}$$

where t_i is the t-value related to the i-th model parameter, v_{ii} is the ii-th term of the parametric variance-covariance matrix, $t(\cdot)$ is the t-value distribution with a $[1/2 + (1$ $\alpha)/2]\%$ confidence level and (N_{sp} - N_θ) degrees of freedom, where α is the statistical level of significance. As can be inferred from the previous equation, each t_i value increases at each sampling time both for the information acquired by the sample itself in terms of v_{ii} and for the variation of the t-distribution due to the stepwise increase of N_{sp} at each sampling point.

A statistically sound parameter estimation is reached for all the model parameters if each parametric t-value is greater than the $(1\ \alpha)\%$ confidence reference t-value ($t^{ref}_{1-\alpha}$), which is defined as:

$$t^{ref}_{1-\alpha} = t(1-\alpha, N_{sp} - N_{\theta}) \quad . \tag{A.8}$$

The IDRO method consists therefore in conducting the experiment in order to overcome the $t^{ref}_{1-\alpha}$-value threshold by the end of experiment, through the online redesign of the experiment. The objective function has been defined as the time integral of the difference between the updating time for each redesign and the total duration of the experiment.[2] Moreover, the optimization is constrained by imposing that, for each redesign task, the predicted precision of all parameters should be greater than a preset threshold.

The mathematical expression of the optimal values of each $\boldsymbol{\varphi}_j$ component of the $\boldsymbol{\varphi}$ design vector, as defined in Eqs. A.2 and A.4, is therefore:

$$\begin{aligned} & \boldsymbol{\varphi}^{opt}_j = \underset{\boldsymbol{\varphi}_j}{\operatorname{argmin}} \int_{t^{up,j}}^{\tau} (t - t^{up,j})dt \\ \text{s.t.} & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , \\ & \mathbf{C} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{E} \left\{ \max_{i=1,\dots,N_{sp}} \left[w_{j-1} \cdot t^{ref}_{1-\alpha}|_{t=t^{up,j}} - t_i|_{t=t^{up,j}} \right] \right\} = 0 \end{aligned} \tag{A.9}$$

where t is the time variable, $t^{up,j}$ is the j-th updating time, w_{j-1} is a weight corrector that corresponds to the minimum fraction of the $t^{ref}_{1-\alpha}$ to be reached by each parametric t-value by the end of the j-th updating time and $E(\cdot)$ is the expected value of the function in brackets, evaluated in the parameter domain $\boldsymbol{\Theta}$. In practice, for each sub-experiment, $\boldsymbol{\varphi}^{opt}_j$ corresponds to the design vector that allows minimizing both a specific percentage of deviation between the $t^{ref}_{1-\alpha}$ and the minimum t_i, and the value of the time $t^{up,j}$ necessary to achieve this result.

The robustness of IDRO is based on the fact that each $\boldsymbol{\varphi}^{opt}_j$ is calculated by evaluating the constraint defined in Eq. A.9, which is an average value obtained through a stochastic simulation with N_{sc} scenarios. Each scenario represents the model response at different $\boldsymbol{\theta}$ values, which are randomly sampled in the parameter domain $\boldsymbol{\Theta}$ of the parameter variance-covariance matrix by assuming a normal distribution $N(\hat{\theta}_i,\ \hat{\sigma_{\theta_i}})$, where $\hat{\theta}_i$ is the estimate of the i-th parameter and $\hat{\sigma_{\theta_i}}$ is the standard deviation of $\hat{\theta}_i$. In order to follow the logic of an A-optimal MBDoE, only the diagonal elements $\hat{\sigma_{\theta_i}}^2$ of the variance-covariance matrix of model parameters are considered during the sampling procedure (i.e. covariance elements are ignored). Although this is an approximation, it has been verified that this choice does not affect the quality of the sampling in a significant way.

The parametric set for each scenario has been randomly generated by considering only the diagonal terms of the parametric variance-covariance matrix (diagonal approach). This simplification is justified by the fact that the average and the standard deviation values of

---

[2] The integral formulation has been chosen in order to amplify the effect of the time distance between $t^{up,j}$ and τ. Alternative formulations are possible.

the sample distribution obtained through the diagonal approach only differ by 1% and 10%, with respect to the case where all the terms of the variance-covariance matrix are included. Furthermore, it should be noted that $\hat{\theta}_i$ and $\hat{\sigma_{\theta_i}}$ values are updated at each $t^{up,j}$ in order to take into account the results obtained through the parameter estimations conducted at the end of each sub-experiment, and the sampling effect is noticeable only at the very beginning of the design procedure (i.e. in the first design).

The logic of IDRO iterative procedure is illustrated in Fig. A.2. The picture simulates the hypothetical behavior of IDRO approach for a two-parameters model. The red solid stepwise profile represents the weighted reference t-value profile, whose value decreases after each sampling point has been collected (initially the profile is constant to a preset value since no distribution can be computed for (N_{sp} - N_θ) < 0). The black solid and dashed stepwise lines represent the time profiles of the t-values for the two model parameters. These profiles will increase after each sampling point is collected as soon as additional information on parametric variability is acquired. Let us assume that in this case the dashed line is related to the most difficult parameter to be estimated (i.e. the parameter with minimum t-value); after fixing a weight w_0 for the first sub-experiment, the IDRO optimization (see Eq. A.9) will minimize the time required for the minimum t-value to become larger than the target $w_0 \cdot t^{ref}_{1-\alpha}$value; the optimal updating time $t^{up,1}$ will be evaluated accordingly.
Note that weight w_0 is initially introduced to set a minimum level of information to be reached within the first update; as will be shown in the actual algorithm, this weight is automatically adjusted to guarantee that the first update does not exploit all the available experimental capability in terms of variations on the input variables (i.e. the maximum number of switches on $\mathbf{u}(t)$, which is set a priori and typically depends on the equipment characteristics) as well as in terms of samples (i.e. the maximum number of measurements that can betaken, which is also set a priori and depends on the experimental settings and on the cost and complexity of measurements). In other words, the procedure is built is such a way that at least one update is always carried out and that the remaining part of the experiment can still be designed according to an optimum criterion.

The experiment then starts and will be run till $t = t^{up,1}$, time at which a parameter estimation is carried out and a new (larger) weight w_1 for $t^{ref}_{1-\alpha}$ is chosen. From this point on, an iterative procedure is implemented until the end of the experiment is reached (in terms of time τ) or $t^{ref}_{1-\alpha}$ is achieved (i.e. $w_j \geq 1$) or there is no experimental design capacity in terms of measurements or input switches. The detailed IDRO procedure (see Fig. A.3) is therefore based on the following steps:

1. get prior knowledge on the parametric set value (and its related uncertainty $\mathbf{\Sigma}_\theta$) and set $j = 1$, where j is the counter index of the sub-experiment to be designed;

2. set w_0, i.e. fix the minimum amount of information, that has to be reached by all the t-values (t_i) at the end of the first sub-experiment, defined by $w_0 \cdot t^{ref}_{1-\alpha}$;

3. design the first sub-experiment from Eq. A.9;

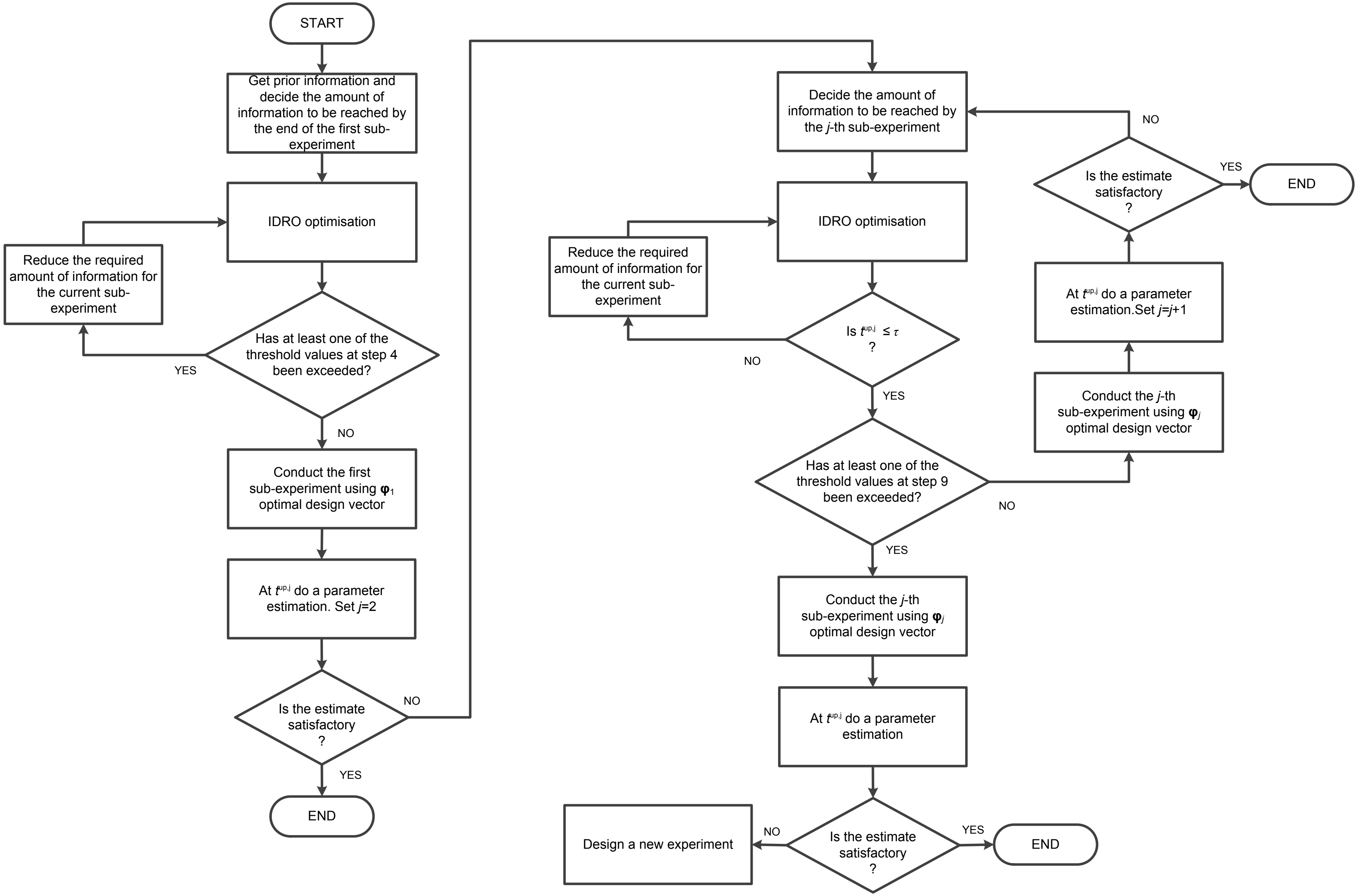


Figure A.2: *Flowchart of IDRO procedure*

4. check if $t^{up,j} < \beta{\cdot}\tau$, $N_{sw,j} < \gamma{\cdot}N_{sw}$ or $N_{sp,j} < N_{sp,j-1} + \delta{\cdot}N_{sp}$ are satisfied[3]; β, γ and δ values are a priori set by the user in order to guarantee both sufficient time and a sensible number of switching levels and sampling points for the following sub-experiment(s);
    (a) if all conditions are satisfied, carry out the sub-experiment, estimate the model parameters and go to step 5.
    (b) otherwise, set $w_0 = k_\varepsilon{\cdot}w_0$ and go to step 3. k_ε represents a preset reduction coefficient that allows to decrease the initial value assumed for w_0 during the iterative procedure;

At the end of the first sub-experiment, the following iterative procedure is used:

5. set $j = j{+}1$;
6. set $w_j = k_w{\cdot}w_{j-1}$, where k_w is the incremental coefficient. This index, whose value is preset by the user, represents the relative difference between the optimal weight w_{j-1} used for the design of the (j-1)-th sub-experiment and the initial tentative weight w_j tested at the beginning of the optimization loop for the j-th sub-experiment;
7. design the first sub-experiment from Eq. A.9;
8. check if $t^{up,j} \leq \tau$;
    (a) if true, go to step 9;
    (b) otherwise, set $w_j = k_w{\cdot}w_j$ and go to step 7;
9. run the j-th sub-experiment (till $t^{up,j}$) and estimate the model parameters;
10. if $w_{j-1} < 1$ and/or $t^{up,j} < t^{up,j-1} + \beta(\tau - t^{up,j-1})$ and/or $N_{swj} < \gamma(N_{sw} - N_{sw,j-1})$ and/or $N_{spj} < \delta(N_{sp} - N_{sp,j-1})$:
    (a) then go to step 5;
    (b) otherwise, stop;
11. if a satisfactory parameter estimation is not achieved, design a new experiment.

Two simulated case studies were examined in order to compare, in terms of parametric estimation precision at the end of the experimental run, standard MBDoE and OMBRE

---

[3] As anticipated above, this condition is required to ensure that one update is carried out and still the remaining part of the experiment can be designed in an optimal way. In other words, we do not want that: (a) the update is scheduled too close to the end of the experiment; (b) there is no further possibility to excite the system (no enough switches on input variables); (c) there is no possibility to collect a reasonable additional number of samples. Note that parameters β, γ and δ are the only ones that are indeed set by the users and somewhat reflects his/her knowledge on the experimental facility and the viability of performing portions of experiment.

approaches with the proposed IDRO methodology; the two case studies differ in terms of number of measured responses, simulated experimental error on measured variables, and for the absence/presence of active constraints on the state variables. For standard MBDoE and OMBRE designs, the A-criterion is used because of its similarity with the IDRO approach, which is based on the diagonal terms of the variance-covariance matrix. This will allow a more fair comparison between different criteria.

The gPROMS^{TM} software has been adopted for design and simulation purposes, as well as to conduct the 'in silico' experiments, and to analyze the results coming from them through its *Parameter Estimation* entity. The NLPSQP solver has been used for the optimization/design task; it employs a sequential quadratic approach (SQP) method for the solution of a nonlinear programming (NLP) problem. The gPROMS^{TM} software is used in conjunction with external subroutines written in Fortran^{TM} via *Foreign Object Interface* in order to implement the inversion of the information matrix and calculate the dynamic t-value profiles required by the IDRO approach.

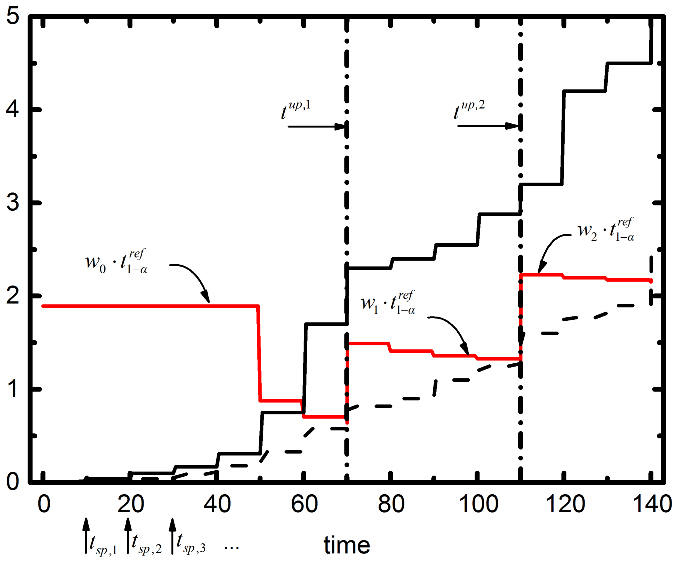


Figure A.3: *The logic of IDRO approach applied to a generic two-parameters model: the red solid stepwise profile represents the dynamic weighted* $t^{ref}_{1-\alpha}$ *profile (at each* $t^{up,j}$ *the weight* w_{j-1} *is modified, as described in Fig. A.3). The black solid and dashed stepwise profiles represent the parametric t-values of the two model parameters, whose values increase with the number of collected samples. The dash-dot vertical lines represent the updating times* $t^{up,j}$ *for the experimental run.*

A.5 Case study 1: fermentation bioreactor

Conventional MBDoE, OMBRE and IDRO approaches are here compared and applied to a model that simulates a biomass fermentation process (Espie and Macchietto (1989); Munack and Posten (1989)). The model consists of the following set of equations:

$$\begin{aligned}
&\frac{dx_1}{dt} = (r_g - u_1 - \theta_4) \cdot x_1 \\
&\frac{dx_2}{dt} = -\frac{r_g x_1}{\theta_3} + u_1(u_2 - x_2) \\
&r_g = \frac{\theta_1 x_2}{(\theta_2 + x_2)} \\
&\hat{y}_i = x_i, \quad i = 1, 2
\end{aligned} \quad , \tag{A.10}$$

where x_1 is the biomass concentration (g/L), x_2 is the substrate concentration (g/L), u_1 is the dilution factor (h^{-1}), and u_2 is the substrate concentration in the feed (g/L). The achievable conditions that characterize the experimental tests are the dilution factor u_1 (range 0.05-0.20 h^{-1}) and the substrate concentration in the feed u_2 (range 5-35 g/L). These manipulated inputs are approximated by piecewise constant profiles over $N_{sw} = 6$ switching intervals. The initial biomass concentration x_1^0 is allowed to assume values in the range 1-10 g/L, whereas the substrate concentration x_2^0 is set to 0 g/L, respectively. It is assumed that both x_1 and x_2 can be measured during the experiment (i.e. $\hat{\mathbf{y}} = [x_1 \; x_2]^{\mathrm{T}}$) and that its total duration is $\tau = 48$ h. Measurements are affected by Gaussian noise with zero mean and variance-covariance matrix given by:

$$\boldsymbol{\Sigma}_{\mathbf{y}} = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.05 \end{pmatrix} . \tag{A.11}$$

The ‘real’ system is assumed to be characterized through the parameter set $\boldsymbol{\theta} = [0.310 \; 0.180 \; 0.550 \; 0.050]^{\mathrm{T}}$. It is then assumed that the initial parameter estimates are affected by a 90% relative error with respect to the true parameter values (in this case, $\boldsymbol{\theta}^0 = [0.589 \; 0.018 \; 1.045 \; 0.005]^{\mathrm{T}}$). The design and parameter estimation activities are implemented on the normalized parametric set $\boldsymbol{\Theta}$, defined as the ratio between the current parametric estimates and the true parameter values $\boldsymbol{\theta}^0$. Note that this normalization procedure is an ideal one since in practice the true value of the parameters is unknown. However, in a simulated case study, the approach is useful to have an immediate grasp of the accuracy of the estimates (in fact, it is also useful for numerical reasons; in real applications the initial guess of the parameter values can be used for normalization).

The following design configurations have been considered and compared:

- Instance I: conventional MBDoE;
- Instance II: OMBRE. It is assumed that $n^{up} = 1$ update may be performed and that for each manipulated input three switching levels are allowed in each updating interval; the length of eac hredesign time window has been fixed equal to 24 h;
- Instance III: IDRO. In this case the iterative procedure, described in detail in Section A.4, requires to set some preliminary constants. In particular, the threshold constants β, γ and δ are set to 2/3; the reduction coefficient k_ε set to 0.9, while the

incremental coefficient k_w set equal to 8. The number of scenarios N_{sc} implemented in the stochastic simulation is set to 25: this value is a good compromise between the high computational cost required by the optimization task and the necessity to simulate a reliable representation of model parameters distribution (using a computer with CPU 2.40 GHz and RAM 8GB, IDRO requires about 5 h of calculation whereas OMBRE needs 45 min; this time is required for the first design and, in fact, the calculation could be carried out before initiating the experiment; the redesign calculation time is about 1 h). It is assumed that measurements ($N_{sp} = 8$) are equidistantly distributed along the maximum experiment duration τ. Results are discussed in the following sections for the three design configurations in terms of manipulated inputs, simulated profiles and *a-posteriori* statistics on the final parameter estimation.

Table A.1: *Case study 1 - Fermentation model: A-optimal MBDoE, OMBRE and IDRO. Results from final parameter estimation including standard deviation and t-value 95% of the normalized parametric set. Double asterisks (**) denote t-values failing the t-test (i.e. the t-value is smaller than the reference one).*

		Normalized parameter	Real value	Final value	Standard deviation	t-value 95%
Conventional MBDoE		Θ_1	1	1.016	0.017	28.17
		Θ_2	1	0.544	0.325	**0.769****
		Θ_3	1	1.046	0.025	18.94
		Θ_4	1	1.188	0.082	5.075
					Reference t-value 95%	1.783
OMBRE	1^{st} sub-exp	Θ_1	1	2.304	17.61	**0.047****
		Θ_2	1	3.136	32.70	**0.035****
		Θ_3	1	1.013	0.037	9.925
		Θ_4	1	1.028	0.095	3.905
					Reference t-value 95%	2.132
	2^{nd} sub-exp	Θ_1	1	2.528	20.42	**0.057****
		Θ_2	1	2.742	27.77	**0.045****
		Θ_3	1	0.988	0.034	13.19
		Θ_4	1	0.948	0.088	4.92
					Reference t-value 95%	1.783
IDRO	1^{st} sub-exp	Θ_1	1	1.042	0.043	9.959
		Θ_2	1	0.799	0.435	**0.731****
		Θ_3	1	1.053	0.060	7.196
		Θ_4	1	1.242	0.213	2.382
					Reference t-value 95%	1.943
	2^{nd} sub-exp	Θ_1	1	0.999	0.013	34.72
		Θ_2	1	0.912	0.145	2.88
		Θ_3	1	1.003	0.023	20.29
		Θ_4	1	1.020	0.079	5.917
					Reference t-value 95%	1.783

A.5.1 Instance 1: conventional MBDoE

Results in terms of predicted profiles and manipulated inputs are given in Figs. A.4(a) and A.4(d). Table A.1 shows that the results obtained in terms of parameter precision are unsatisfactory; in fact, it is not possible to achieve a statistically sound estimation of parameter θ_2.

A.5.2 Instance 2: OMBRE

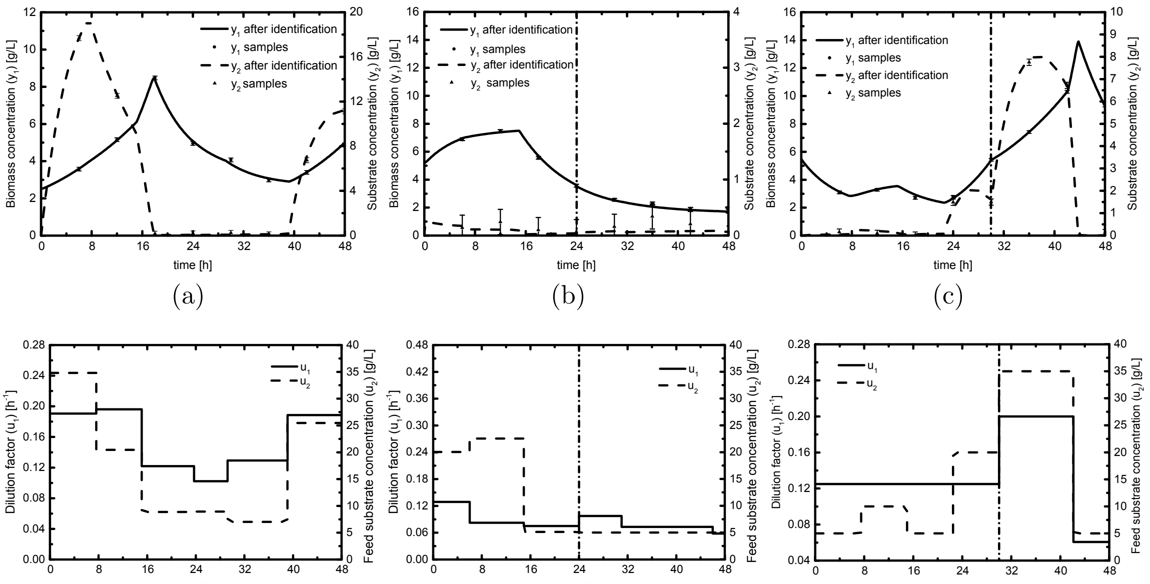
Results in terms of predicted profiles and manipulated inputs are given in Figs. A.4(b) and A.4(e). Also in this case, the estimation of the parameters values is quite unsatisfactory: as summarized in Table A.1 parameters θ_2 and θ_3 cannot be estimated in a statistically satisfactory way by the end of the experimental run. Note that OMBRE is outperformed by the conventional MBDoE configuration. The initial uncertainty on model parameters is such that during the first sub-experiment it is not possible to gather sufficient information to produce an effective redesign. In fact, this issue was highlighted in the original article by the authors themselves (Galvanin et al. (2009a)) when they observed that in the case of poor initial information the quality of the experiments is highly affected by the updating policy and that scarcely informative redesigns could be obtained in the initial phases.

A.5.3 Instance 3: IDRO

Results in terms of predicted profiles and manipulated inputs are given in Figs. A.4(c) and A.4(f), while the parameter estimates are given in A.1. It is clear that, compared to MBDoE and OMBRE configuration, IDRO gives the final best estimates in terms of accuracy, thanks both to the robust approach and the innovative info-based updating time optimization which guarantees the attainment of a minimum level of information in the initial experimental phases. In fact, this is the only case where a satisfactory estimation of all parameters can be achieved within the duration on one experiment. Note that in some cases also a robust formulation for OMBRE may produce similarly good results, but the performance is very much dependent on the updating policy, which is dealt explicitly with only in the IDRO approach.

A.6 Case study 2: a physiological model for type 1 diabets mellitus

This second case study considers a model of glucose homeostasis for the simulation of type 1 diabetes mellitus (Galvanin et al. (2011)), based on a previous work (Lynch and Bequette (2002)). The model is described by the following set of differential and algebraic equations:



(a) (b) (c)

(d) (e) (f)

Figure A.4: *Case study 1 - Fermentation model: results in terms of predicted profiles for biomass (y_1) and substrate (y_2) concentration for: (a) A-optimal MBDoE, (b) OMBRE, (c) IDRO. The experimental samples are indicated by circles (y_1) and triangles (y_2) with error bars. Optimal profiles for the dilution factor (u_1) and the substrate concentration in the feed (u_2) are reported for: (d) A-optimal MBDoE, (e) OMBRE, (f) IDRO.*

$$\begin{aligned}
&\frac{dC_g}{dt} = \theta_1 C_g - X_I(C_g - C_{g,b}) + \frac{D_{meal}(t)}{V_G} \\
&\frac{dX_I}{dt} = \theta_2 X_I + \theta_3 I_c, \quad \frac{dI_c}{dt} = -\theta_4(I_c - I_{c,b}) + \frac{u(t)}{V_I} \\
&\frac{dC_{gsc}}{dt} = \frac{C_g - C_{gsc}}{5} - R_{ut} \\
&D_{meal}(t) = \frac{5A_g t e^{-0.05t}}{2}, \quad \hat{y} = C_{gsc,tot} = C_{gsc} + C_{gsc,b}
\end{aligned}, \tag{A.12}$$

where C_g is the blood glucose concentration (mg/dL), X_I the variation of insulin concentration (mU/L) in the inaccessible compartment with respect to basal value, I_c the variation of insulin concentration (mU/L) with respect to basal value, u(t) the rate of infusion of exogenous insulin (mU/min). The measured response is the total subcutaneous glucose concentration $C_{gsc,tot}$, modeled by a first-order 5 min lag between C_g and C_{gsc}. The meal disturbances $D_{meal}(t)$ is expressed through Hovorka et al. (2004) correlation, with A_g being the amount of carbohydrates of the meal, here set to be 30 gCHO (fixed). The constant basal parameter are set to the following values: the basal glucose concentrations in the blood $C_{g,b}$ and $C_{gsc,b}$ are assumed to be equal to 81 mg/dL, the basal insulin concentration $I_{c,b}$ is 15 mU/L, the glucose distribution volume V_G is 120 dL, the insulin distribution volume V_I is 12 L and the tissue rate utilization R_{ut} is 0.75 mg/dL/min. With respect to the original formulation only the inequality constraint:

$$-y + G_d \leq 0 \tag{A.13}$$

is considered, where $G_d = 70$ mg/dL is the lower conservative threshold on the subcutaneous glucose concentration (y). This bound is a hard constraint not to be violated because it prevents hypoglycemic conditions ($C_{gsc,tot} < 76$ mg/dL) during real tests. The achievable condition that characterize the experimental tests is the dilution factor $u(t)$ (range 0 115 mU/min); this manipulated input is approximated as a piecewise constant function with $N_{sw} = 8$ switching intervals. It is assumed that $N_{sp} = 14$ samples (equally distributed) can be taken during a $\tau = 480$ min long experiment. Measurements are affected by Gaussian noise with zero mean and a standard deviation of $\sigma = 5$ mg/dL on y. The real system is represented by the parameter set $\boldsymbol{\theta} = [0.017\ 0.032\ 1.540\text{E-}5\ 0.096]^{\mathrm{T}}$; it is then assumed that the initial parameter estimates are affected by a 70% relative error with respect to the true parameter values (in this case, $\boldsymbol{\theta}^0$=$[0.029\ 0.010\ 2.620\text{E-}5\ 0.027]^{\mathrm{T}}$). The design and parameter estimation activities are based on the normalized parametric set $\boldsymbol{\Theta}$, defined through the ratio between the current parametric estimates and the true parameter values. Also in this case study, conventional MBDoE, OMBRE and IDRO approaches have been compared:

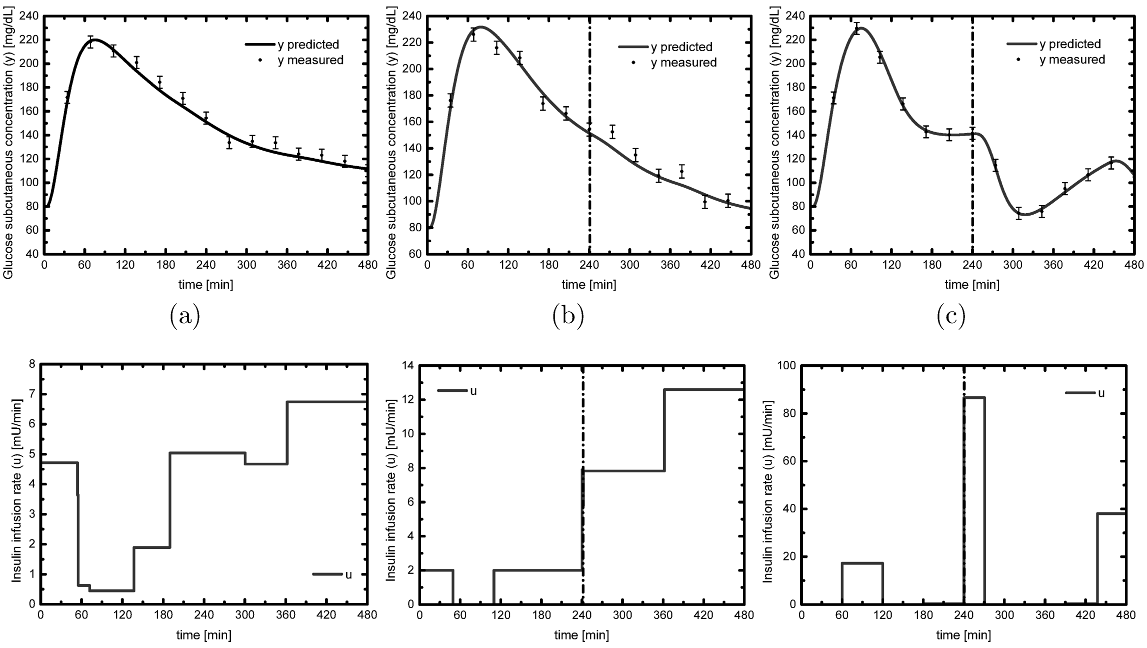
- Instance I: conventional MBDoE;
- Instance II: OMBRE. It is assumed that $n^{up} = 1$ update may be performed and that for each manipulated input three switching levels are allowed in each updating interval; the length of each redesign time window has been fixed equal to 240 min.
- Instance III: IDRO. In the same way described for the first case study, the threshold constants β, γ and δ are set to 2/3; the reduction coefficient k_ε set to 0.9, while the incremental coefficient k_w set equal to 8. The number of scenarios N_{sc} implemented in the stochastic simulation is set to 25. The calculation time is similar to the one required for case study 1.

A.6.1 Instance 1: conventional MBDoE

Results in terms of predicted profiles and manipulated inputs are given in Figs. A.5(a) and A.5(d). The parameter estimation task proves to be totally unsatisfactory; in fact, it is not possible to achieve an accurate estimation of parameters θ_1, θ_2 and θ_3 (see Table A.2).

A.6.2 Instance 2: OMBRE

Results in terms of predicted profiles and manipulated inputs are given in Figs. A.5(b) and A.5(e). The OMBRE approach allows for a slightly better performance (see Table A.2), but still both parameters θ_2 and θ_3 cannot be estimated in a statistical satisfactory way.



(a) (b) (c)

(d) (e) (f)

Figure A.5: *Case study 2 - Diabetes model: results in terms of predicted profiles for biomass (y_1) and substrate (y_2) concentration for: (a) A-optimal MBDoE, (b) OMBRE, (c) IDRO. The experimental samples are indicated by circles (y_1) and triangles (y_2) with error bars. Optimal profiles for the dilution factor (u_1) and the substrate concentration in the feed (u_2) are reported for: (d) A-optimal MBDoE, (e) OMBRE, (f) IDRO.*

Once again, the behavior reflects the initial sub-optimal design due to the initial parametric mismatch and a non-robust approach for the optimal design/redesign of the experiment.

A.6.3 Instance 3: IDRO

Results in terms of predicted profiles and manipulated inputs are given in Figs. A.5(c) and A.5(f). Table A.2 shows that in this case the final results obtained in terms of parameter accuracy are greatly improved and all parameters pass the t-test. Note that IDRO sets the updating time exactly after 240 min like IDRO (compare Figs. A.5(b) and A.5(c)); however, the first sub-experiment is more informative as it is explicitly designed to guarantee some level of information in a robust manner. That cannot be ensured in the OMBRE case, whose efficiency is affected by the initial parametric mismatch.

A.7 Conclusions

An innovative approach for a structured online model based redesign of the experiment (IDRO) has been presented in this paper. This technique allows for the determination of the optimal updating points at which exploiting the experimental information for the online redesigning procedure, thanks to both a robust approach and an info-driven objective function to be optimized in the design phase. This novel redesign criterion can be adopted

Table A.2: *Case study 2 - Diabetes model: A-optimal MBDoE, OMBRE and IDRO. Results from final parameter estimation including standard deviation and t-value 95% of the normalized parametric set. Double asterisks (**) denote t-values failing the t-test (i.e. the t-value is smaller than the reference one).*

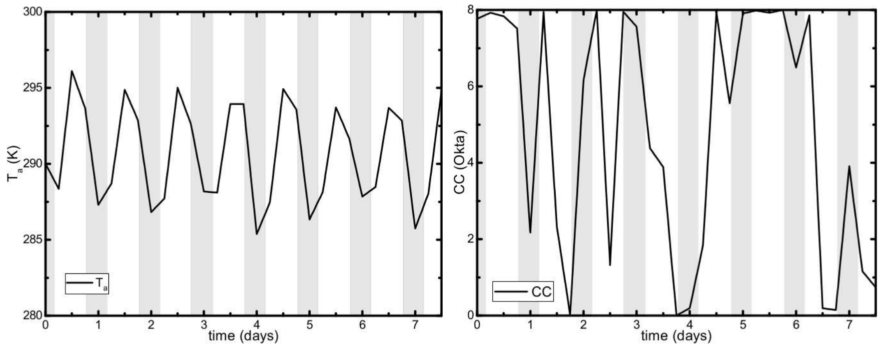
		Normalized parameter	Real value	Final value	Standard deviation	t-value 95%
Conventional MBDoE		Θ_1	1	1.127	0.057	**1.660****
		Θ_2	1	1.358	1.190	**0.741****
		Θ_3	1	1.648	0.037	**0.474****
		Θ_4	1	1.019	0.082	2.955
					Reference t-value 95%	1.812
OMBRE	1^{st} sub-exp	Θ_1	1	0.654	0.260	**0.793****
		Θ_2	1	0.586	0.678	**0.272****
		Θ_3	1	0.388	0.565	**0.216****
		Θ_4	1	0.312	0.920	**0.107****
					Reference t-value 95%	2.354
	2^{nd} sub-exp	Θ_1	1	1.019	0.185	2.478
		Θ_2	1	1.657	0.560	**1.327****
		Θ_3	1	1.586	0.848	**0.840****
		Θ_4	1	1.067	0.073	3.352
					Reference t-value 95%	1.812
IDRO	1^{st} sub-exp	Θ_1	1	1.059	0.116	2.873
		Θ_2	1	1.339	0.582	**0.723****
		Θ_3	1	1.436	0.522	**0.865****
		Θ_4	1	1.044	0.256	**1.282****
					Reference t-value 95%	2.354
	2^{nd} sub-exp	Θ_1	1	1.106	0.055	9.035
		Θ_2	1	1.034	0.170	2.736
		Θ_3	1	1.247	0.169	3.322
		Θ_4	1	1.082	0.073	6.625
					Reference t-value 95%	1.812

to increase the reliability of the design task; a specific advantage over the conventional design techniques for improving parameter estimation is given by the fact that information is exploited during the experimental run; moreover, with respect to online redesign techniques (such as OMBRE), the great improvement consists in treating the design in a robust way by optimizing both the redesign structure and the available information in a synergistic way. Two distinct simulated case studies have been used to assess the effectiveness of the new technique and compare it with conventional redesign approaches: one relates to a bioreactor system, while the other one to a physiological system describing the effect of insulin on patients suffering from diabetes; results of both case studies show the higher efficiency of the proposed technique to exploit the information coming from the running experiment in an info-driven way. Although the computational burden is still very demanding, much of the effort is required during the first calculation only, i.e. before starting the experiment. Furthermore, more tailored optimization methods and the usage of parallel computing (especially for stochastic simulations) could improve the algorithm efficiency drastically.

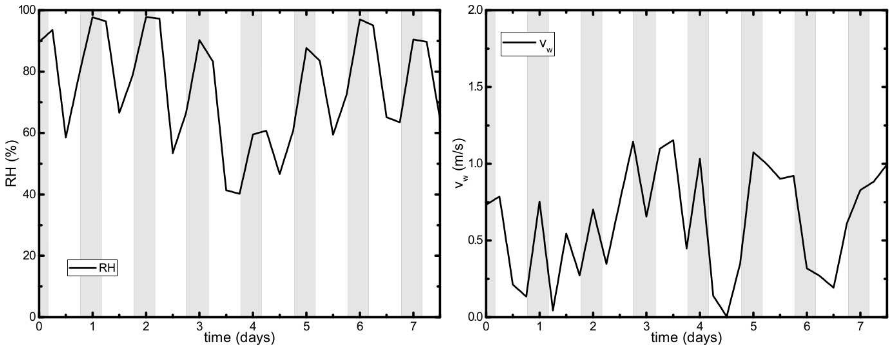
Appendix B

Meteorological data profiles

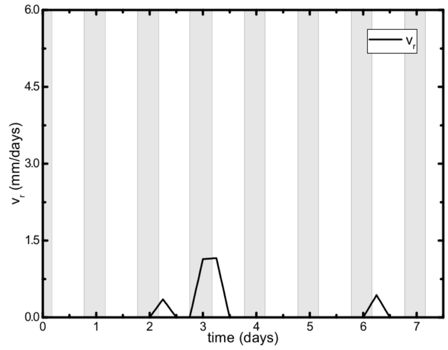
The weather data extracted from the European Centre for Medium-Range Weather Forecast (ECMWF) website were: the air temperature T_a, the sky cloudiness CC, the relative humidity RH, the wind velocity v_w and the rain volumetric flux v_r. Figures B.1, B.2, B.3 and B.4 represent, respectively, the meteorologic data used to simulate summer (June-July), winter and spring 2012.



(a) June: Air temperature (b) June: Cloudiness

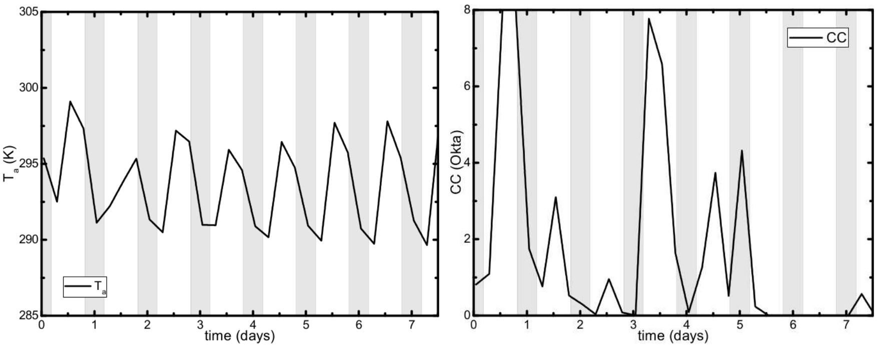


(c) June: Relative humidity (d) June: Wind velocity

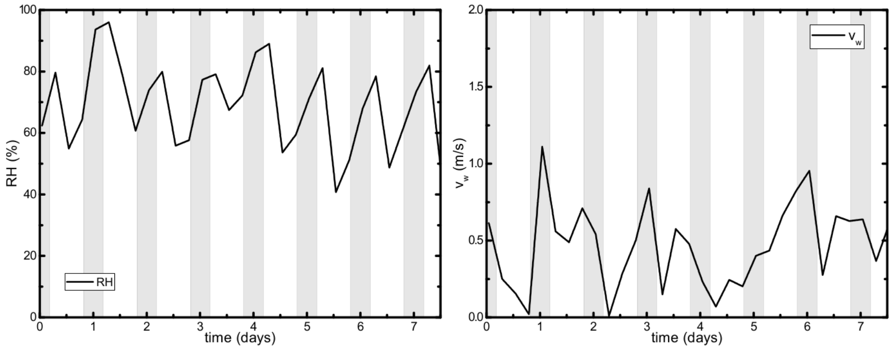


(e) June: Rain velocity

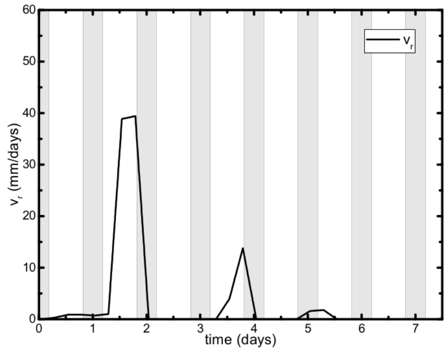
Figure B.1: *Meteorological data: Summer case-June (The background is colored in white at daytime and in grey at nighttime.)*



(a) July: Air temperature (b) July: Cloudiness

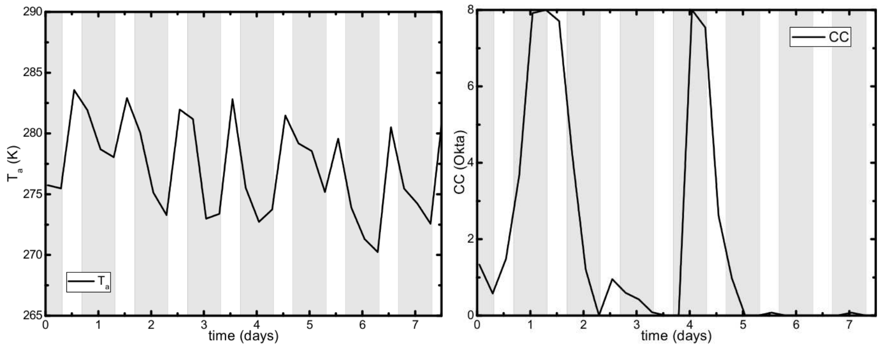


(c) July: Relative humidity (d) July: Wind velocity

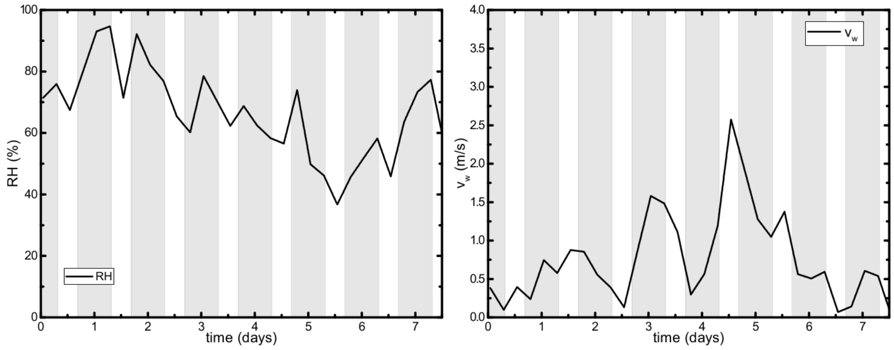


(e) July: Rain velocity

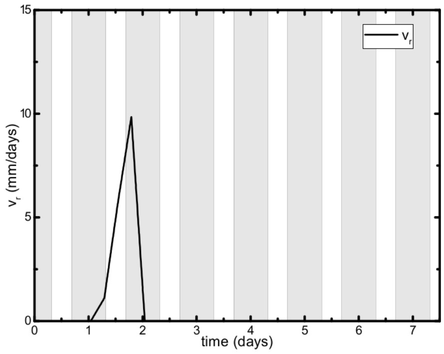
Figure B.2: *Meteorological data: Summer case-July (The background is colored in white at daytime and in grey at nighttime.)*



(a) Winter: Air temperature (b) Winter: Cloudiness

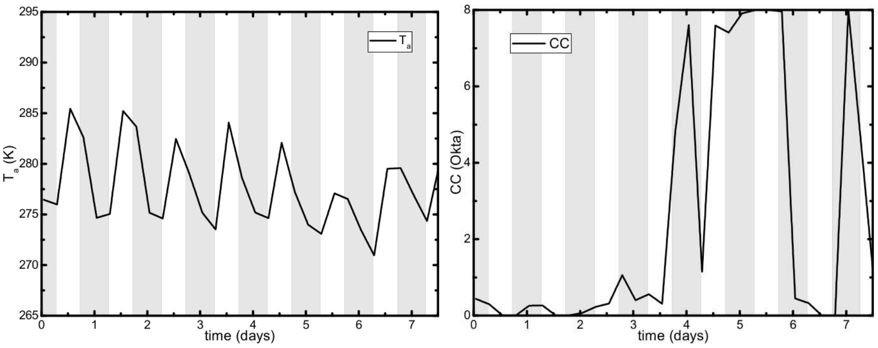


(c) Winter: Relative humidity (d) Winter: Wind velocity

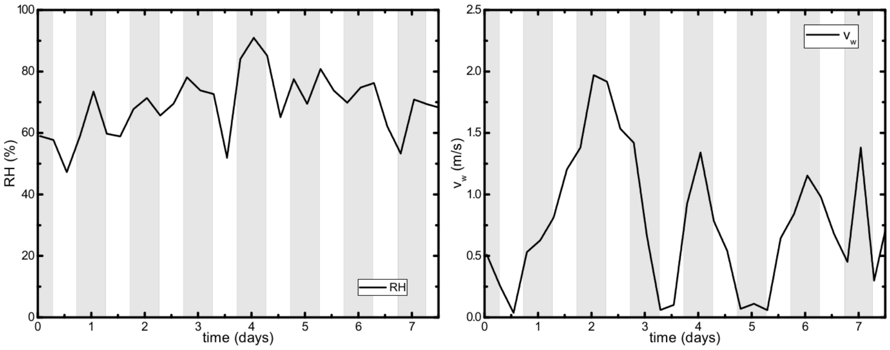


(e) Winter: Rain velocity

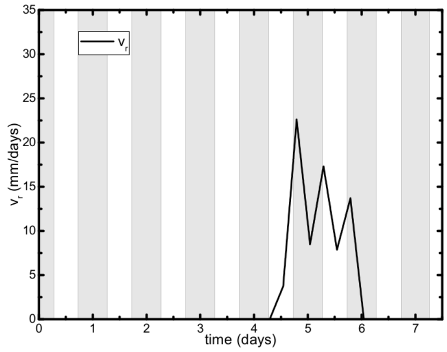
Figure B.3: *Meteorological data: Winter case (The background is colored in white at daytime and in grey at nighttime.)*



(a) Spring: Air temperature (b) Spring: Cloudiness



(c) Spring: Relative humidity (d) Spring: Wind velocity



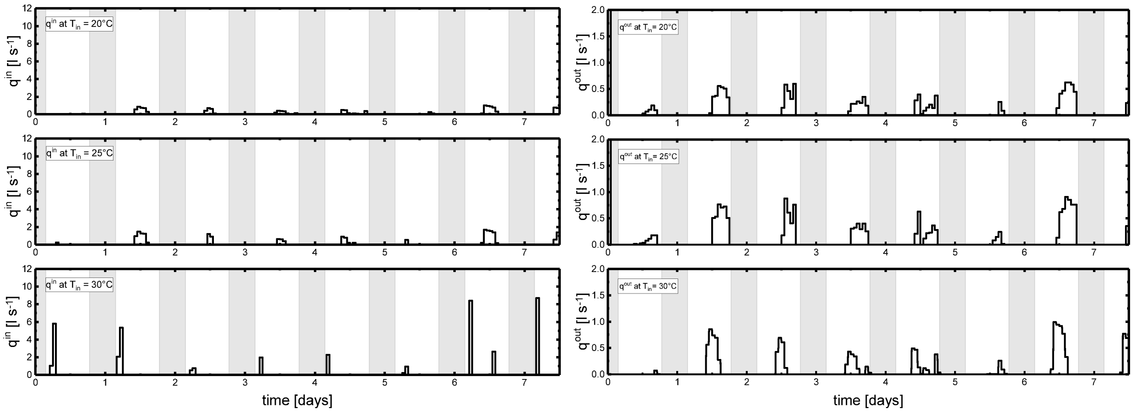
(e) Spring: Rain velocity

Figure B.4: *Meteorological data: Spring case (The background is colored in white at daytime and in grey at nighttime.)*

Appendix C

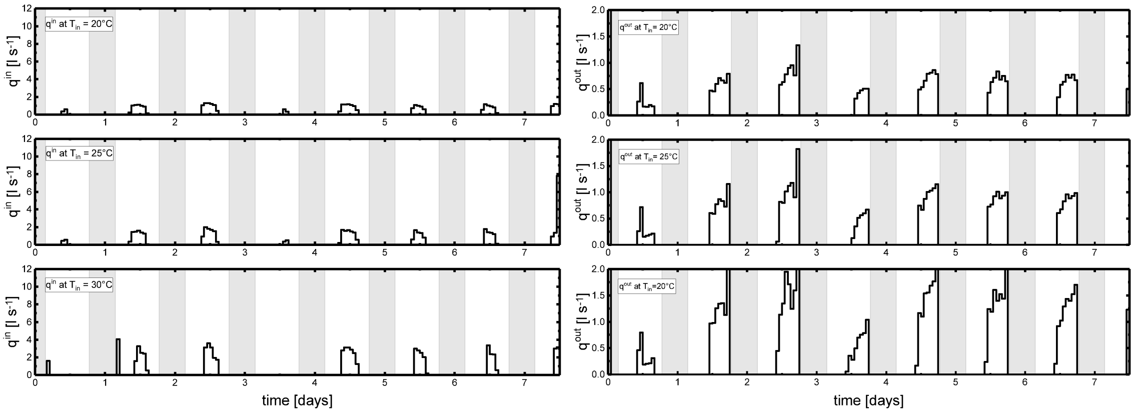
Optimal control dynamics graphs: Ch.5 extension

In this Appendix we show all the graphical representation of q^{in} and q^{out} optimal dynamics obtained for the base case (Figures C.1 and C.2), the 3σ case (Figures C.3 and C.4) and the conservative case (Fig. C.5). The optimal profiles, obtained at different T_{in} values, are reported for all the weather conditions tested in Chapter 5.



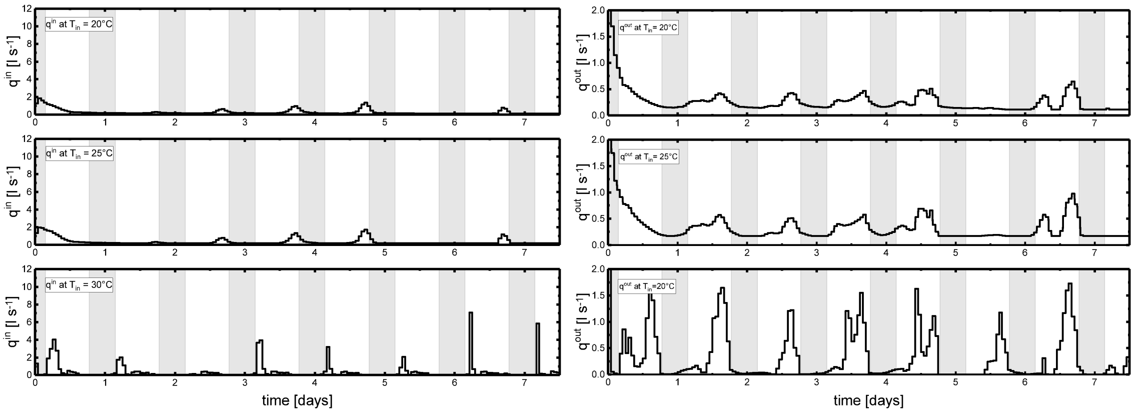
(a) (b)

Figure C.1: *Base case: optimal (a) q^{in} and (b) q^{out} profiles in the month of June at T_{in} = 20, 25, 30 °C (the background is colored in white at daytime and in grey at nighttime).*



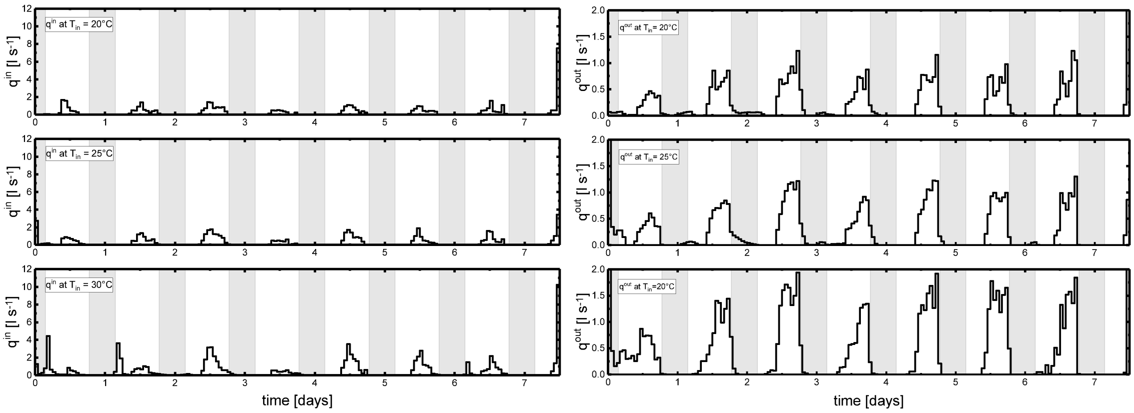
(a) (b)

Figure C.2: *Base case: optimal (a) q^{in} and (b) q^{out} profiles in the month of July at T_{in} = 20, 25, 30 ° C (the background is colored in white at daytime and in grey at nighttime).*



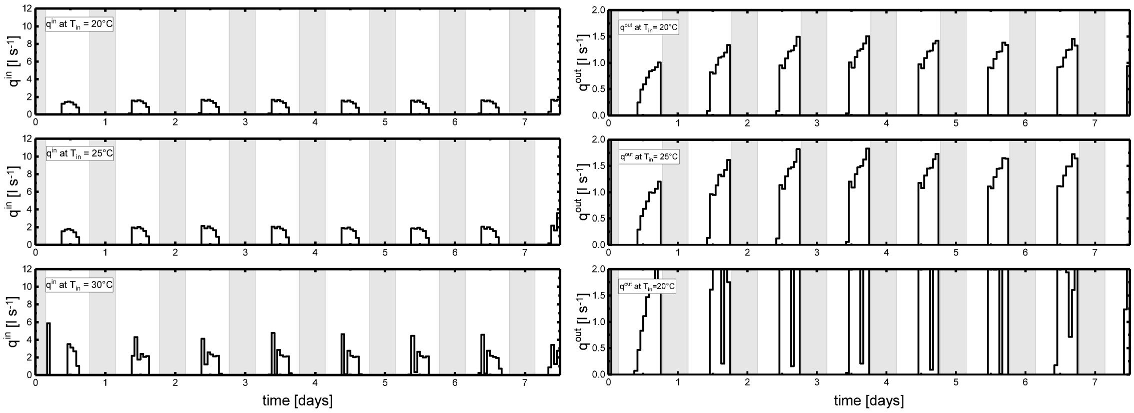
(a) (b)

Figure C.3: *3σ case: optimal (a) q^{in} and (b) q^{out} profiles in the month of June at T_{in} = 20, 25, 30 ° C (the background is colored in white at daytime and in grey at nighttime).*



(a) (b)

Figure C.4: *3σ case: optimal (a) q^{in} and (b) q^{out} profiles in the month of July at T_{in} = 20, 25, 30 ° C (the background is colored in white at daytime and in grey at nighttime).*



(a) (b)

Figure C.5: *Consevative case: optimal (a) q^{in} and (b) q^{out} profiles in the month of July at T_{in} = 20, 25, 30 °C (the background is colored in white at daytime and in grey at nighttime).*

Bibliography

Ali, H., Cheema, T. A., and Park, C. W. (2017). Numerical prediction of heat transfer characteristics based on monthly temperature gradient in algal open raceway ponds. *Biotechnology and Bioengineering*, 106:7–17.

Antal, T. K., Kovalenko, I. B., Rubin, A. B., and Tyystjärvi, E. (2013). Photosynthesis-related quantities for education and modelling. *Photosynthesis research*, 117(1):1–30.

Araya, B., Nobre, B., Reis, A., Chamy, R., and Poirrier, P. (2014). Evaluation of the simultaneous production of lutein and lipids using a vertical alveolar panel bioreactor for three *Chlorella* species. *Algal Research*, 6:218–222.

Aro, E.-M. (2016). From first generation biofuels to adavanced solar biofuels. *Ambio*, 45:24–31.

Asprey, S. P. and Macchietto, S. (2002). Designing robust optimal dynamic experiments. *Journal of Process Control*, 24:545–556.

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

Bard, Y. (1974). *Noninear parameter estimation.* New York USA: Academic Press Inc.

Baroukh, C., Muñoz-Tamayo, R., J.P., S., and Bernard, O. (2014). Drum: A new framework for metabolic modeling under non-balanced growth. application to the carbon metabolism of unicellular microalgae. *PLoS ONE*, 9(8):1–15.

Barz, T., López Cárdenas, D. C., Arellano-Garcia, H., and Wozny, G. (2013). Experimental evaluation of an approach to online redesign of experiments for parameter determination. *AIChe Journal*, 59:1981–1995.

Béchet, Q., Chambonnière, P., Shilton, A., and Guizard, G. (2015a). Algal productivity modeling: a step toward accurate assessments of full-scale algal cultivation. *Biotechnology and Bioengineering*, 112(5):987–996.

Béchet, Q., Chambonnière, P., Shilton, A., and Guizard, G. (2015b). Algal productivity modeling: a step toward accurate assessments of full-scale algal cultivation. *Biotechnology and Bioengineering*, 112(5):987–996.

Béchet, Q., Shilton, A., and 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.

Béchet, Q., Shilton, A., Park, J. B. K., Craggs, R. J., and Guieysse, B. (2011). Universal temperature model for shallow algal ponds provides improved accuracy. *Environmental Science & Technology*, 45:3702–3709.

Belay, A. (1997). *Mass culture of Spirulina outdoors - the Earthrise Farms experience*. Taylor & Francis, London, UK.

Bennion, E., Ginosar, D., Moses, J., F, A., and JC, Q. (2015). Lifecycle assessment of microalgae to biofuel: comparison of theromochemical processing pathways. *Applied Energy*, 154:1062–1071.

Berenguel, M., Rodrıguez, F., Acién, F., and Garcıa, J. (2004). Model predictive control of ph in tubular photobioreactors. *Journal of Process Control*, 14(4):377–387.

Bernard, O. (2011). Hurdles and challenges for modelling and control of microalgae for co2 mitigation and biofuel production. *Journal of Process Control*, 21:1378–1389.

Bernard, O., Boulanger, A. C., Bristeau, M. O., and Saint-Marie, J. (2013). A 2d model for hydrodynamics and biology coupling applied to algae growth simulations. *ESAIM: Mathematical and Numerical Analysis*, 47(5):1387–1412.

Bernard, O., Mairet, F., and Chachuat, B. (2015). Modelling of microalgae culture systems with applications to control and optimization. In *Microalgae Biotechnology*, pages 59–87. Springer.

Bernard, O. and Remond, B. (2012). Validation of a simple model accounting for light and temperature effect on microalgae growth. *Bioresource technology*, 123:520–527.

Bernardi, A., Nikolaou, A., Meneghesso, A., Morosinotto, T., Chachuat, B., and Bezzo, F. (2016). High-fidelity modelling methodology of light-limited photosynthetic production in microalgae. *PLoS ONE*, 11(4):e015238.

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

Bona, F., Capuzzo, A., Franchino, M., and Maffei, M. E. (2014). Semicontinuous nitrogen limitation as convenient operation strategy to maximize fatty acid production in *Neochloris oleoabundans*. *Algal Research*, 5:1–6.

Borowitzka, M. A. (1999). Commercial production of microalgae: ponds, tanks, tubes and fermenters. *Journal of Biotechnology*, 70:313–321.

Borowitzka, M. A. (2005). *Culturing microalgae in outdoor ponds.* Elsevier (2005).

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.

Camacho, E. F. and Alba, C. B. (2013). *Model predictive control.* Springer Science & Business Media.

Camacho-Rubio, F., Camacho, F. G., Sevilla, J. M. F., Chisti, Y., and Grima, E. M. (2003). A mechanistic model of algal photosynthesis in microalgae. *Biotechnology and Bioengineering*, 81:459–473.

Chakrabarty, A., Buzzard, G. T., and Rundell, A. I. (2013). Model-based design of experiments for cellular processes. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 5(2):181–203.

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

Douglas, J. (1988). *Conceptual design of chemical processes.* McGraw-Hill.

Droop, M. R. (1968). Vitamin b12 and marine ecology. iv. the kinetics of uptake growth and inhibition in *Monochrysis lutheri*. *Journal of the Marine Biological Association of the United Kingdom*, 48(3):689–733.

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

Duffie, J. A. and Beckman, W. A. (1958). *Solar Engineering of Thermal Processes.* Little, Brown & Co., Boston.

Dugdale, R. C. (1967). Nutrient limitation in the sea: dynamics, identification and significance. *Lymnology and Oceanography*, 12:685–695.

Ebenhöh, O., Fucile, G., Finazzi, G., Rochaix, J.-D., and Goldschmidt-Clermont, M. (2014). Short-term acclimation of the photosynthetic electron transfer chain to changing light: a mathematical model. *Philosophical transactions of the Royal Society B*, 369(1640):20130223.

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

EIA (2016). International energy outlook 2016. World total primary energy consumption by region. `https://www.eia.gov/outlooks/aeo/data/browser/`. Available on line.

Espie, D. and Macchietto, S. (1989). The optimal design of dynamic experiments. *AIChe Journal*, 35:223–229.

Falkowski, P. G. and Raven, J. A. (2007). *Acquatic Photosynthesis, 2nd Edition*. Princeton University Press.

Flynn, K. (1991). A mechanistic model for describing dynamic multi-nutrient, light, temperature interactions in phytoplankton. *Ecological Modelling*, 23:977–997.

Foley, P., Beach, E., and Zimmerman, J. (2011a). Algae as source of renewable chemicals: opportunities and challenges. *Green Chemistry*, 13(6):1399–1405.

Foley, P. M., Beach, E. S., and Zimmerman, J. B. (2011b). Algae as a source of renewable chemicals: opportunities and challenges. *Green Chemistry*, 13(6):1399–1405.

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

Galvanin, F., Barolo, M., and Bezzo, F. (2008). Towards on-line model-based design of experiments. *Computer Aided Chemical Engineering*, 25:349–354.

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

Galvanin, F., Barolo, M., Pannocchia, G., and Bezzo, F. (2012). Online model-based redesign of experiments with erratic models: a disturbance estimation approach. *Computers & Chemical Engineering*, 42:138–151.

Galvanin, F., Boschiero, A., Barolo, M., and Bezzo, F. (2011). Model-based design of experiments in the presence of continuous measurements systems. *Industrial and Engineering Chemistry Research*, 50:2167–2175.

Galvanin, F., Macchietto, S., Barolo, M., and Bezzo, F. (2009b). Optimal design of clinical tests for the identification of physiological models of type 1 diabetes mellitus. *Industrial and Engineering Chemistry Research*, 48:1989–2002.

Gárcia-Camacho, F., Sánchez-Mirón, A., Grima, E. M., Camacho, F. G., and Merchuck, J. C. (2012). A mechanistic model of photosynthesis in microalgae including photoacclimation dynamics. *Journal of Theoretical Biology*, 304:1–15.

Geider, R., MacIntyre, H., and Kana, T. (1998). A dynamic regulatory model of phytoplankton acclimation to light, nutrient and temperature. *Lymnology and Oceanography*, 43:679–694.

Gerencsér, L., Hjalmarsson, H., and Mårtensson (2009). Identification of ARX systems with non-stationary inputs asymptotic analysis with application to adaptive input design. *Automatica*, 45(3):623–633.

Guerrero-Lemus, R. and Martinez-Duart, J. M. (2013). *Renewable Energies and CO2. Lecture Notes in Energy.* Springer-Verlag (London).

Haiden, T., Janousek, M., Bauer, P., Bidlot, J., Dahoui, M., Ferranti, L., Prates, F., Richardson, D. S., and Vitart, F. (2015). *Evaluation of ECMWF forecasts, including 2014-2015 upgrades.* European Centre for Medium Range Weather Forecasts, Shinfield Park, England.

Han, B. P. (2002). A mechanistic model of algal photoinhibition induced by photodamage to photosystem-ii. *Journal of Theoretical Biology*, 214:519–527.

Hartmann, F., Demory, D., Combe, C., Hamouda, R., Boulanger, A. C., Bristeau, M. O., Saint-Marie, J., Sialve, B., P., S. J., Rabouille, S., Sciandra, A., and Bernard, O. (2014). Growth rate estimation of algae in raceway ponds: A novel approach. In *9th World Congress The International Federation of Automatic Control (Cape Town, South Africa. August 24-29, 2014)*, pages 6216–6221.

Havlik, I., Lindler, P., Sheper, T., and Reardon, K. F. (2013). Online monitoring of large cultivations of microalgae and cyanobacteria. *Trends in Biotechnology*, 31(7):406–414.

Houska, B., Telen, D., Logist, F., Diehl, M., and Van Impe, J. F. M. (2015). An economic objective for the optimal design of nonlinear dynamic processes. *Automatica*, 51:98–103.

Hovorka, R., Canonico, V., Chassin, L. J., Haueter, U., Massi-Benedetti, M., Orsini Federici, M., Pieber, T. R., Shaller, H. C., Schaupp, L., Vering, T., and Wilinska, M. E. (2004). Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement*, 25(4):905–920.

Hussain, A., Arif, S. M., and Aslam, M. (2017). Emerging renewable and sustainable energy technologies. *Renewable and sustainable energy reviews*, 71:12–28.

IEA (2016). Key world energy statistics 2016. `https://www.iea.org/publications/freepublications/publication/key-world-energy-statistics.html`. Available on line.

Jiménez, C., Cossío, B. R., Labella, D., and Niell, F. X. (2003). The feasibility of industrial production of *Spirulina* (arthrosphira) in southern spain. *Aquaculture*, 217:179–190.

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

Kiran, B., Kumar, R., and Deshmukh, D. (2014). Perspectives of microalgal biofuels as a renewable source of energy. *Energy conversion and management*, 88:1228–1244.

Klass, D. L. (1998). *Biomass for Renewable Energy, Fuels, and Chemicals.* Academic Press, San Diego (1998).

Kookos, I. and Perkins, J. (2004). *The integration of Process Design and Control.* Elsevier BV.

Körkel, S., Kostina, E., Bock, H. G., and Schlöder, J. P. (2004). Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optimization methods and software*, 19:327–338.

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

Lee, E., Jalalizadeh, M., and Zhang, Q. (2015). Growth kinetic models for microalgae cultivation: a review. *Algal Research*, 12:497–512.

Lindqvist, K. and Hjalmarssonn, H. (2001). Identification for control: adaptative input design using convex optimization. In *In: Proceedings of the 40th IEEE Conference on Decision and Control (Orlando, Florida USA, December 2001)*, pages 4326–4331.

Lorenz, E., Hurka, J., Heinemann, D., and Beyer, H. (2009a). Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. *IEEE Journal of selected topics in applied earth observation and remote sensing*, 2(1):2–10.

Lorenz, E., Remund, J., Müller, S., Traunmüller, W., Steinmaurer, G., Pozo, D., Ruiz-Arias, J., Fanego, V., Ramirez, L., Romeo, M., Kurz, C., Pomares, L., and Guerrero, C. (2009b). Benchmarking of different approaches to forecast solar irradiance. *Proceedings of the 24th European Photovoltaic Solar Energy Conference, 21-25 September 2009, Hamburg, German*, pages 4199–4208.

Lynch, S. M. and Bequette, B. W. (2002). Model predictive control of blood glucose in type i dia-betics using subcutaneous glucose measurements. In *In: Proceedings of the American Control Conference (Anchorage, AK May 2002)*, pages 4039–4043.

Mairet, F., Bernard, O., Masci, P., Lacourt, T., and Sciandra, A. (2011). Modelling neutral lipid production by the microalga *Isochrysis* aff. *galbana* under nitrogen limitation. *Bioresource Technology*, 102:142–149.

Marthews, T. R., Malhi, Y., and Iwata, H. (2012). Calculating downward longwave radiation under clear and cloudy conditions over a tropical lowland forest site: an evaluation of model schemes for hourly data. *Theoretical and applied climatology*, 107(3-4):461–477.

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

Medipally, S. R., Yusoff, F. M., Banerjee, S., and Shariff, M. (2015). Microalgae as sustainable renewable energy feedstock for biofuel production. *BioMed Research International*, 2015:1–13.

Mehra, R. K. (1974). Optimal input signals for parameter estimation in dynamic systems: survey and new results. *IEEE Transactions on Automatic Control*, 19:753–768.

Meinshausen, M., Meinshausen, N., Hare, W., Raper, S. C. B., Frieler, K., Knutti, R., Frame, D. J., and Allen, M. R. (2009). Greenhouse-gas emission targets for limiting global warming to 2 °C. *Nature*, 458:1158–1162.

Mendoza, J., Granados, M., Godos, I. D., Acién, F., Molina, E., Banks, C., and Heaven, S. (2013). Fluid-dynamic characterization of real-scale raceway reactors for microalgae production. *Biomass and Bioenergy*, 54:267–275.

Meneghesso, A., Simionato, D., Gerotto, C., Rocca, N. L., Finazzi, G., and Morosinotto, T. (2016). Photoacclimation of photosynthesis in the *Eustigmatophycean Nannochloropsisgaditana*. *Photosynthesis research*, 129(3):291–305.

Mohr, S. H., Wang, J., Ellem, G., Ward, J., and Giurco, D. (2015). Projection of world fossil fuels by country. *Fuel*, 141:120–135.

Molina Grima, E., Belarbi, E.-H., Acién Fernandéz, F., Robles Medina, A., and Chisti, Y. (2003). Recovery of microalgal biomass and metabolites: process options and economics. *Biotechnology Advances*, 20:491–515.

Moody, J. W., McGinty, C. M., and Quinn, J. C. (2014). Global evaluation of biofuel potential from microalgae. *Proceedings of the National Academy of Sciences*, 111(23):8691–8696.

Moreno-Garcia, L., Adjallé, K., Barnabé, S., and Raghavan, G. S. V. (2017). Microalgae biomass production for a biorefinery system: Recent advances and the way towards sustainability. *Renewable and sustainable energy reviews*, 76:493–506.

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

Munack, A. and Posten, C. (1989). Design of optimal dynamical experiments for parameter estimation. In *In: Proceedings of the American Control Conference (Pittsburgh, USA May 1989)*, pages 2011–2016.

Muñoz-Tamayo, R., Mairet, F., and Bernard, O. (2013). Optimizing microalgal production in raceways systems. *Biotechnology Progress*, 29 (2):543–552.

Nashawi, I. S., Malallah, A., and Al-Bisharah, M. (2009). Forecasting world crude oil production using multicyclic Hubbert model. *Energy and Fuels*, 24(3):1788–1800.

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

Norsker, N. H., Barbosa, M. J., Vermuë, M. H., and Wijffels, R. H. (2011). Microalgal production - a close look at the economics. *Biotechnology Advances*, 29:24–27.

Oldewurtel, F., Parisio, A., Jones, C. N., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B., and Morari, M. (2012). Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45:15–27.

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.

Pann, Y. (2011). A large and persistent carbon sink in the world's forest. *Science*, 333:988–993.

Papadakis, I. A., Kotzabakis, K., and Lika, K. (2012). Modeling the dynamic modulation of light energy in photosynthetic algae. *Journal of Theoretical Biology*, 300:254–264.

Park, S. and Li, Y. (2015). Integration of biological kinetics and computational fluid dynamics to model the growth of *Nannochloropsis salina* in an open channel raceway. *Biotechnology and Bioengineering*, 112(5):923–933.

Patel, A., Gami, B., Patel, P., and Patel, B. (2017). Microalgae: Antiquity to era of integrated technology. *Renewable and sustainable energy sources*, 71:535–547.

Peeters, J. C. H. and Eilers, P. (1978). The relationship between light intensity and photosyntheis in natural assemblage of marine phytoplankton. *Hydrobiological Bullettin*, 12:134–136.

Perez, R., Lorenz, E., Pelland, S., Beauharnois, M., Van Knowe, G., Hemker Jr, K., Heinemann, D., Remund, J., Müller, S., Traunmüller, W., Pozo, D., Ruiz-Arias, J., Lara-Fanego, V., Ramirez-Santinosa, L., Gaston-Romero, M., and Pomares, L. (2013). Comparison of numerical weather prediction solar irradiance forecasts in the us, canada and europe. *Solar energy*, 94:305–326.

Platt, T., Gallegos, C. L., and Harrison, W. G. (1980). Photoinhibition of photosynthesis in natural assemblages of marine phytoplankton. *Journal of Marine Research*, 38:687–701.

Prasad, V. and Vlachos, D. G. (2008). Multiscale model and informatics-based optimal design of experiments: application to the catalytic decomposition of ammonia on ruthenium. *Industrial and Engineering Chemistry Research*, 47:6555–6567.

Pukelsheim, R. (1993). *Optimal design of experiments.* New York USA: J. Wiley & Sons.

Pulz, O. (2001). Photobioreactors: production systems for phototrophic microorganisms. *Applied Microbiology and Biotechnology*, 57(3):287–293.

Ramos Tercero, E., Domenicali, G., and Bertucco, A. (2014). Autotrophic production of biodiesel from microalgae: an updated process and economic analysis. *Energy*, 76:807–815.

Richardson, J. W., Johnson, M. D., and Outlaw, J. L. (2012). Economic comparison of open pond raceways to photo bio-reactors for profitable production of algae for transportation fuels in the southwest. *Algal Research*, 1:93–100.

Ridolfi, L., Zittelli, G. C., and Bassi, N. (2009). Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnology and Bioengineering*, 102(1):100–112.

Riley, J. A. (1946). Factors controlling phytoplankton populations on georges bank. *Journal of Marine Research*, 6:54–63.

Rochaix, J. D. (2011). Reprint of: Regulation of photosynthetic electron transport. *Biochima and Biophysica acta*, 1807:878–886.

Rogers, J. N., Rosenberg, J. N., Guzman, B. J., Oh, V. H., Mimbela, L. E., Ghassemi, A., Betenbaugh, M. J., Oyler, G. A., and Donohue, M. D. (2014). A critical analysis of paddlewheel-driven raceway ponds for algal biofuel production at commercial scales. *Algal research*, 4:76–88.

Roser, M. and Ortiz-Ospina, E. (2017). World population growth. `https://ourworldindata.org/world-population-growth/`. Available on line.

Schenkendorf, R. and Mangold, M. (2013). Online model selection approach based on unscented Kalman filtering. *Journal of Process Control*, 23:44–57.

Shafiee, S. and Topal, E. (2009). When will fossil fuel reserves be diminished? *Energy Policy*, 37(1):181–189.

Sharma, J., Najafi, M., and Qasim, S. (2013). Preliminary cost estimation models for construction, operation, and maintenance of water treatment plants. *Journal of Infrastructure Systems*, 19(4):451–464.

Skjånes, K., Rebours, C., and Lindblad, P. (2013). Potential for green microalgae to produce hydrogen, pharmaceuticals and other high value products in a combined process. *Critical reviews in biotechnology*, 33(2):172–215.

Slegers, P., Van Beveren, P., Wijffels, R., Van Straten, G., and Van Boxtel, A. (2013). Scenario analysis of large scale algae production in tubular photobioreactors. *Applied energy*, 105:395–406.

Smil, V. (2008). *Energy in Nature and Society: general energetics of complex systems.* MIT Press.

Solimeno, A., Samsó, R., Uggetti, E., Sialve, B., Steyer, J.-P., Gabarró, A., and García, J. (2015). New mechanistic model to simulate microalgae growth. *Algal Research*, 12:350–358.

Steele, J. H. (1962). Environmental control of photosynthesis in the sea. *Lymnology and Oceanography*, 7:137–150.

Stigter, J. D., Vries, D., and Keesman, K. J. (2006). On adaptive optimal input design: a bioreactor casestudy. *AIChe Journal*, 52:3290–3296.

Takache, H., Christophe, G., Cornet, J. F., and Pruvost, J. (2010). Experimental and theoretical assessment of maximum productivities for the microalgae *Chlamydomonas reinhardtii* in two different geometries of photobioreactors. *Biotechnology progress*, 26:431–440.

Tebbani, S., Lopes, F., Filali, R., Dumur, D., and Pareau, D. (2014a). *CO2 biofixation bymicroalgae: modeling, estimation and control.* Wiley ISTE.

Tebbani, S., Lopes, F., Filali, R., Dumur, D., and Pareau, D. (2014b). Nonlinear predictive control for maximization of co2 bio-fixation by microalgae in a photobioreactor. *Bioprocess and biosystems engineering*, 37(1):83–97.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., Siqueira, M. F. D., Granger, A., Hannan, L., Hughes, L., Huntley, B., Jaarsveld, A. S. V., Midgley, G. F., Miles, L., Huerta, M. A. O., Peterson, A. T., Phillips, O. L., and Williams, S. E. (2004). Extinction risk from climate change. *Nature*, 427(6970):145–148.

Tikhonov, A. N. and Vershubskii, A. V. (2014). Computer modeling of electron and proton transport in chloroplasts. *Biosystems*, 121:1–21.

Vollenweider, R. A. (1965). Calculation models of photosynthesis-depth curves and some implications regarding day rate estimates in primary production measurements. *Memorie dell'Istituto italiano di idrobiologia*, 18:425–457.

Vredenberg, W. J. and Bulychev, A. A. (2010). Photoelectrochemica control of the balance between cyclic- and linear electron transport in photosystem 1.Algorithm for p700+ induction kinetics. *Biochimica et Biophysica acta*, 1797:1521–1532.

White, R. L. and Ryan, R. A. (2015). Long-term cultivation of algae in open-raceway ponds: Lessons from the field. *Industrial Biotechnology*, 11(4):213–220.

Wigmosta, M. S., Coleman, A. M., Skaggs, R. J., Huesemann, M. H., and Lane, L. J. (2011). National microalgae biofuel production potential and resource demand. *Water Resources Research*, 47(3).

Wu, X. and Merchuck, J. C. (2001). A model integrating fluid dynamics in photosynthesis and photoinhibition processes. *Chemical Engineering Science*, 56(11):3527–3538.

Yang, J., Xu, M., Zhang, X., Hu, Q., Sommerfeld, M., and Chen, Y. (2011). Life-cycle analysis on biodiesel production from microalgae: water footprint and nutrients balance. *Bioresource technology*, 102(1):159–165.

Zong, Y., Kullmann, D., Thavlov, A., Gehrke, O., and Bindner, H. W. (2012). Application of model predictive control for active load management in a distributed power system with high wind penetration. *IEEE Transactions on Smart Grid*, 3(2):1055–1062.

Zonneveld, C. (1998). A cell-based model for the chlorophyll a to carbon ratio in phytoplankton. *Ecological Modelling*, 113:55–70.

Zullo, L. (1991). Computer aided design of experiments. an engineering approach. In *Ph.D. Thesis – UK University of London.*

Acknowledgements

"Invero è vita, con la sua magia..."

There are many people I would like to thank for their technical and emotional support during these three (...more or less) years of research. First of all I would like to thank my supervisor, Prof. Fabrizio Bezzo, for both his help and support for this Thesis and the encouragement to become a writer in case of failure as chemical engineer.
Special thanks to Dr. Andrea Bernardi for his invaluable help and for the precious and helpful discussions made during the final part of my Thesis work. Thanks to Prof. Tomas Morosinotto and Dr. Andrea Meneghesso for their patience: it is not so easy to explain the basics of biotechnology to a mere engineer in such a rapid way!
I am also grateful to Prof. Olivier Bernard and Dr. Quentin Béchet for their support during my visit at INRIA-Biocore (France). Thanks a lot for your great contribution to my Thesis work and for the wonderful months spent in Sophia-Antipolis in 2015.
Thanks to all the past and present members of CAPE-Lab, with whom I shared this journey: in particular to Prof. Massimiliano Barolo, Dr. Federico Galvanin, Dr. Pierantonio Facco, Natascia Meneghetti, Filippo Dal Pastro, Ricardo Ortiz, Martina Largoni, Myriam Ferrari, Christopher Castaldello, Gabriele Bano, Mario Soprana and Federico d'Amore.
Warm thanks to all the master students I had the pleasure to work with during my PhD: Thomas Paladin, Luigi Bosa and Martino Trabuio; it has been a source of pride and gratification to have worked with all of you.

Un enorme grazie ai Reparti di Neurochirurgia, di Radioterapia Oncologica e di Oncologia dell'Ospedale Cà Foncello di Treviso, senza il quale il sottoscritto avrebbe trovato qualche difficoltà in più a scrivere queste parole. Un grazie particolare al Dr. Stefano Lamon, per il suo interesse nel seguire la storia di questa Tesi durante i nostri incontri mensili e per i suoi suggerimenti sul raccogliere alghe in quel di Jesolo durante il periodo estivo.

E infine grazie alla mia famiglia, che dopo quasi trent'anni ha ancora la pazienza di starmi vicino di fronte a ogni piccola o grande difficoltà, e ai miei amici, che dopo quasi trent'anni lo sono ancora.