

Latvijas Lauksaimniecības universitāte  
Informācijas Tehnoloģiju fakultāte

Latvia University of Agriculture  
Faculty of Information Technologies

**Mg.sc.ing. Ivars MOZGA**

## **BIOĶĪMISKO TĪKLU STACIONĀRO STĀVOKĻU OPTIMIZĀCIJAS PROCEDŪRA**

## **STEADY STATE OPTIMIZATION PROCEDURE OF BIOCHEMICAL NETWORKS**

Promocijas darba  
**KOPSAVILKUMS**  
doktora grāda ieguvei  
Informācijas tehnoloģiju nozarē (Dr.sc.ing.)

**SUMMARY**  
of the thesis for the acquisition of a doctoral degree  
in the field of Information Technologies (Dr.sc.ing.)

**Ivars Mozga** \_\_\_\_\_  
Paraksts

**Jelgava, 2012**

## INFORMĀCIJA

**Darba izpildes vieta** – Latvijas Lauksaimniecības universitāte, Informācijas Tehnoloģiju fakultāte, Datoru sistēmu katedra.

**Promocijas darba zinātniskais vadītājs** – Latvijas Lauksaimniecības universitātes asociētais profesors Dr.sc.ing. Egils Stalidzāns.

**Promocijas darba konsultants** – Latvijas Universitātes Mikrobioloģijas un biotehnoloģijas institūta vadošais pētnieks Dr.habil.biol. Pēteris Zikmanis.

### **Promocijas darba zinātniskā aprobācija noslēguma posmā.**

1. Prezentēts doktoranta atklātajā seminārā 2009. gada 11. jūnijā.
2. Apspriests un aprobēts ESF projekta ietvaros „Latvijas starpaugstskolu zinātniskās grupas izveide sistēmbioloģijā” (līguma Nr. 2009/0207/1DP/1.1.2.0/09/APIA/VIAA/128)
3. Atzīts par pilnībā sagatavotu un pieņemtu LLU Informācijas Tehnoloģiju Promocijas padomē 2012. gada 29. februārī.

**Darbs akceptēts** LLU Informācijas Tehnoloģiju fakultātes Datoru sistēmu katedras paplašinātā akadēmiskā sēdē 2011.gada 25.maijā. Protokols Nr.1.

Promocijas darbs izstrādāts ar **Eiropas Sociālā fonda (ESF)** projekta „Atbalsts doktorantūras studijām un pēcdoktorantūras pētījumiem inženierzinātnēs un mezzinātnē” (līguma Nr. 2004/0004/VPD1/ESF/APIA/04/NP/3.2.3.1/0001/0005/0067) un ESF projekta „Latvijas starpaugstskolu zinātniskās grupas izveide sistēmbioloģijā” (līguma Nr. 2009/0207/1DP/1.1.2.0/09/APIA/VIAA/128) atbalstu.

### **Oficiālie recenzenti:**

1. Latvijas Universitātes asociētais profesors Dr.sc.comp. Juris Vīksna;
2. Rīgas Tehniskās universitātes profesors Dr.habil.sc.ing. Juris Merkurjevs;
3. Abo Akademi (Turku, Somija) universitātes profesors PhD Ion Petre.

**Promocijas darba aizstāvēšana** notiks LLU Informācijas tehnoloģiju nozares promocijas padomes atklātā sēdē 2012. gada 22. maijā Jelgavā, Lielā ielā 2, Informācijas Tehnoloģiju fakultātes 218. auditorijā plkst. 13:00.

Ar promocijas darbu var iepazīties LLU Fundamentālajā bibliotēkā, Lielā ielā 2, Jelgavā un <http://llufb.llu.lv/llu-theses.htm>.

**Atsauksmes sūtīt** Promocijas padomes sekretārei – Lielā ielā 2, Jelgava, LV-3001; tālrunis: 63005621; e-pasts: [tatjana.tabunova@llu.lv](mailto:tatjana.tabunova@llu.lv). Atsauksmes vēlams sūtīt skenētā veidā ar parakstu.

**Padomes sekretāre** – LLU lektore Mg.paed. Tatjana Tabunova.

## SATURA RĀDĪTĀJS

INFORMĀCIJA .....	2
PROMOCIJAS DARBA APROBĀCIJA .....	4
IEVADS .....	7
Tēmas aktualitāte .....	7
Promocijas darba mērķis un uzdevumi .....	8
Pētījuma metodes .....	9
Zinātniskais jauninājums un praktiskā vērtība .....	9
Darba struktūra un apjoms .....	10
1. SISTĒMBIOLOĢIJAS PIEEJA BIOPROCESU VADĪBAS PROBLĒMU RISINĀŠANĀ .....	10
2. BIOĶIMISKO TĪKLU OPTIMIZĀCIJAS UZDEVUMU RISINĀŠANAS PEEJU APSKATS .....	11
3. OPTIMIZĀCIJAS PROCEDŪRAS DARBĪBAS UN TO SECĪBA .....	12
Kritērija un modeļa izvēle .....	13
Optimizācijas potenciāla noteikšana .....	14
Maināmo parametru kombināciju ranžēšana pēc efektivitātes .....	15
Risinājumu pielietojamības analīze .....	17
4. KRITĒRIJA VĒRTĪBAS KONVERĢENCES ANALĪZE STACIONĀRĀ STĀVOKĻA OPTIMIZĀCIJAS EKSPERIMENTA LAIKĀ .....	18
Konverģences analīzes rīks ConvAn .....	19
SECINĀJUMI .....	20
Galvenie darba rezultāti .....	20
Secinājumi un attīstības perspektīvas .....	22
LITERATŪRAS SARAKSTS .....	46

## PROMOCIJAS DARBA APROBĀCIJA

Promocijas darbā veikto pētījumu rezultāti ir atspoguļoti sekojošās publikācijās:

1. Mozga I., Stalidzāns E. (2006) Acquisition of dynamic parameters of biological process from experimental data. International scientific conference “Information Technologies for Rural Development”, pp. 131-140. (Indeksēts AGRIS datu bāzē).
2. Stalidzāns E., Gailums A., Mozga I., (2006) Life cycle of artificial and biological control system. International conference. International scientific conference “Information Technologies for Rural Development”, pp. 164-172. (Indeksēts AGRIS datu bāzē).
3. Grunde-Zeiferts U., Mozga I., Žukova T., Stalidzāns E. (2006) Terapijas modelēšana apvienojot sistēmbioloģijas un automātiskās vadības teorijas metodes. International scientific conference “Animals. Health. Food hygiene”, pp. 70-74. (Indeksēts AGRIS datu bāzē).
4. Stalidzāns E., Mozga I., (2006) Control oriented dynamic modelling of biological object. Lithuanian University of Agriculture, Biometrics and information technologies in agriculture: research and development, pp. 11-14.
5. Mozga I., Stalidzāns E. (2007) Application of dynamic models of glycolysis developing control system. 6th International Scientific Conference „Engineering for Rural Development”. pp. 119-123. (Indeksēts AGRIS datu bāzē).
6. Mozga I., Grunde-Zeiferts U., Sudars S., Stalidzans E. (2007) Life Cycles and Competition in Modelling of Artificial and Biological Control System. The European Simulation and Modelling Conference, pp. 533-536. (Indeksēts AGRIS, EBSCO, Thomson Reuters datu bāzē).
7. Lagzdina L., Mozga I., Stalidzans E. (2008) Algorithm for bioethanol yield optimisation in *Saccharomyces cerevisiae* glycolysis using ‘COPASI. International scientific conference “Applied Information and Communication Technologies”, pp. 95-101. (Indeksēts AGRIS, EBSCO datu bāzē).
8. Mozga I., Kostromins A., Stalidzans E. (2011) Forecast of numerical optimization progress of biochemical networks. 10th International Scientific Conference „Engineering for rural development”, pp. 103-108. (Indeksēts AGRIS, EBSCO, Scopus, Thomson Reuters datu bāzē).
9. Mozga I., Stalidzans E. (2011a) Optimization protocol of biochemical networks for effective collaboration between industry representatives, biologists and modellers. 9th Annual Industrial Simulation Conference “ISC’2011”, pp. 91-96. (Indeksēts Thomson Reuters datu bāzē).
10. Mozga I., Stalidzans E. (2011b) Convergence Dynamics of Biochemical Pathway Steady State Stochastic Global Optimization, 12th IEEE

- International Symposium on Computational Intelligence and Informatics, CINTI 2011 - Proceedings, pp. 231-235. (Indeksēts IEEE Xplore, Scopus datu bāzē).
11. Mozga I., Stalidzans E. (2011c) Convergence Dynamics of Biochemical Models To The Global Optimum, 3rd IEEE International Conference on e-Health and Bioengineering, EHB 2011 – Proceedings, pp. 227-230. (Indeksēts IEEE Xplore, Scopus datu bāzē).
  12. Kostromins A., Mozga I., Stalidzans E. (2012) ConvAn: a convergence analyzing tool for optimization of biochemical networks. Biosystems, 108(1-3), pp. 73-77. Elsevier. doi:10.1016/j.biosystems.2011.12.004. (Indeksēts Scopus, ThomsonReuters datu bāzē).

Pētījumos iegūtie rezultāti tika prezentēti sekojošās konferencēs:

1. Mozga I., Stalidzāns E. Acquisition of dynamic parameters of biological process from experimental data. Jelgava, Latvija. 2006. gada 19.-20. oktobris.
2. Stalidzāns E., Gailums A., Mozga I. Life cycle of artificial and biological control system. Jelgava, Latvija. 2006. gada 19.-20. oktobris.
3. Grunde-Zeiferts U., Mozga I., Žukova T., Stalidzāns E. Terapijas modelēšana apvienojot sistēmbioloģijas un automātiskās vadības teorijas metodes. Jelgava, Latvija. 2006. gada 10. novembris.
4. Stalidzāns E., Mozga I. Control oriented dynamic modelling of biological object. Kauņa, Lietuva. 2006. gada 24.-25. novembris.
5. Mozga I., Acquisition methods of SBML biomodels elements on the Internet. Jelgava, Latvija. 2007. gada 16.-18. maijs.
6. Mozga I., Stalidzāns E. Application of dynamic models of glycolysis developing control system. Jelgava, Latvija. 2007. gada 24.-25. maijs.
7. Mozga I., Grunde-Zeiferts U., Stalidzāns E. Glycolysis control modeling using Systems Biology Markup Language (SBML) software. Pērnava, Igaunija. 2007. gada 26.-30. augusts.
8. Mozga I., Stalidzans E., Metabolic Control Analysis tools in SBML modelling software. NJF seminar „Modelling in Agriculture”, Jelgava, Latvija. 2007. gada 18.-20. oktobris.
9. Mozga I., Grunde-Zeiferts U., Sudars S., Stalidzans E. Life Cycles and Competition in Modelling of Artificial and Biological Control System. The European Simulation and Modelling Conference 2007, St.Julian's, Malta. 2007. gada 22.-24. oktobris.
10. Lagzdina L., Mozga I., Stalidzans E., Algorithm for bioethanol yield optimisation in *Saccharomyces cerevisiae* glycolysis using ‘COPASI. International scientific conference “Applied Information and Communication Technologies”, Jelgava, Latvija, 2008. gada 10.-12. aprīlis.

11. Brusbardis V., Stalidzans E., Mozga I. Biological and artificial control systems as competitors. Summer school on systems biology for medical applications, Tenerife, Spānija, 2008. gada 30. septembris – 2. oktobris.
12. Lagzdina L., Mozga I., Stalidzans E., Algorithm for bioethanol yield optimisation in *Saccharomyces cerevisiae* Glycolysis using Copasi. Summer school on systems biology for medical applications, Tenerife, Spānija, 2008. gada 30. septembris – 2. oktobris.
13. Mozga I., Stalidzans E., Computational optimisation of enzyme concentrations. Systems Biochemistry – Linked Focused Meeting, Jorkas universitāte, Jorka, Lielbritānija, 2010. gada 22.-24. marts.
14. Mozga I., Kostromins A., Stalidzans E., Forecast of numerical optimization progress of biochemical networks. 10th International Scientific Conference „Engineering for rural development”, Jelgava, Latvija, 2011. gada 26.-27. maijs.
15. Mozga I., Stalidzans E., Optimization protocol of biochemical networks for effective collaboration between industry representatives, biologists and modellers. The Industrial Simulation Conference 2011, Venēcija, Itālija, 2011. gada 6.-8. jūnijs.
16. Mozga I., Stalidzans E. Convergence Dynamics of Biochemical Pathway Steady State Stochastic Global Optimization. 12th IEEE International Symposium on Computational Intelligence and Informatics, CINTI 2011, Budapešta, Ungārija, 2011. gada 21.-22. novembris.
17. Mozga I., Stalidzans E. Convergence Dynamics of Biochemical Models To The Global Optimum, 3rd IEEE International Conference on e-Health and Bioengineering, EHB2011, Iasi, Rumānija, 2011. gada 24.-26. novembris.

Pētījumos iegūtie rezultāti tika prezentēti sekojošos semināros:

1. Mozga I. Glikolīzes procesa vispārējs raksturojums, Biosistēmu grupas seminārs, Jelgava, Latvija, 2006. gada 9. oktobris.
2. Mozga I. Glikolīzes procesu simulācijas COPASI programmpaketē, Biosistēmu grupas seminārs, Jelgava, Latvija, 2006. gada 30. oktobris.
3. Mozga I. Dažādu glikolīzes modeļu saīdzinājums, Biosistēmu grupas seminārs, Jelgava, Latvija, 2006. gada 6. novembris.
4. Mozga I. Glikolīzes modeļu saīdzinošā analīze, Biosistēmu grupas seminārs, Jelgava, Latvija, 2006. gada 13. novembris.
5. Mozga I. Bioķīmisko tīklu optimizācijas ātruma un precizitātes prognoze, Biosistēmu grupas seminārs, Jelgava, Latvija, 2009. gada 12. februāris.
6. Mozga I. Bioķīmisko tīklu modelēšana ar CellDesigner, Biosistēmu grupas seminārs, Jelgava, Latvija, 2009. gada 21. oktobris.
7. Mozga I. Sistēmbioloģijas rīks SBtoolbox2 Matlab programmā, Biosistēmu grupas seminārs, Jelgava, Latvija, 2010. gada 3. marts.

8. Mozga I. Optimizācijas kritēriju noteikšana COPASI programmatūrā, Biosistēmu grupas seminārs, Jelgava, Latvija, 2010. gada 3. novembris.
9. Mozga I. Bioķīmisko tīklu optimizācija, kritēriju noteikšana, laika precizitātes un prognozes rīks, Biosistēmu grupas seminārs, Jelgava, Latvija, 2010. gada 1. decembris.

Dalība ar promocijas darba izstrādi saistītos projektos:

1. LZP projekts Nr.08.2178 „Biodegvielas ražošanas racionalizācija izmantojot bioķīmisko reakciju tīklu dinamiskos modeļus” (2008. gads, izpildītājs)
2. „Latvijas starpaugstskolu zinātniskās grupas izveide sistēmbioloģijā” Nr.2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128, (no 2009. - , pētnieks)

Ar promocijas darba rezultātiem saistītie goda nosaukumi un apbalvojumi:

The Young Researcher, 3rd IEEE International Conference on e-Health and Bioengineering, EHB 2011, Iasi, Rumānija, 2011. gada 24.-26. novembris.

## IEVADS

### Tēmas aktualitāte

Biotehnoloģija ir jebkurš dzīvo organismu (vai bioloģisko sistēmu vispār) lietojums, lai sekmētu cilvēku labklājību. Biotehnoloģija tiek izmantota pārtikas ražošanā, atkritumu noārdīšanā, medicīnā un daudzās citās jomās (Reiss, 2005).

Eiropas Savienībā (ES) arvien pieaug interese par biodegvielas izmantošanu Eiropā apkārtējās vides aizsardzības un energoapgādes nodrošināšanas dēļ (Kondili, Kaldellis, 2007). Nemot to vērā, ES ar vairāku Direktīvu (EC 2003) palīdzību izteikti atbalsta biodegvielu izmantošanu.

Par vienu no biotehnoloģijas attīstības priekšnosacījumiem ir izvirzījušās informācijas tehnoloģijas, kuru pielietojumu nosaka vairāki iemesli: lielu datu apjomu glabāšana un apstrāde, līdzību analīze, datizraces uzdevumi un dažādi modelēšanas pielietojumi. Modelēšanas uzdevums galvenokārt ir apkopot zināšanas un pārvērst tās formā, kurā tās var salīdzināt ar eksperimentālajiem datiem, apstiprinot vai izslēdzot kādus iepriekšējus pieņēmumus (Mauch et al., 2001). Modeļi ir izmantojami arī dažādu sistēmas izmaiņu ietekmes prognozēšanai. Šo pieeju plaši izmanto metaboliskajā inženierijā, lai meklētu sistēmas izmaiņas, kuras nepieciešamas tās īpašību uzlabošanai (Keasling, 2010; Sendin et al., 2010).

Šūnu bioķīmisko tīklu optimizācija ir sarežģīts un laikietilpīgs uzdevums, it īpaši pieaugot modeļu izmēriem un variējamo parametru skaitam. Tomēr šūnu bioķīmisko tīklu gadījumā spēkā ir vairākas likumsakarības, kuru izmantošana var vienkāršot risināmo uzdevumu. Uzdevuma risinājumu telpu ierobežot var arī ar datu bāzēs pieejamo informāciju par dažāda veida eksperimentiem. Līdz ar to optimizācijas gaitā ir veicams daudzveidīgu pasākumu komplekss, kura laikā skaitliskās optimizācijas metodes ir tikai viens no instrumentiem, kurš jāpielieto īstajā brīdī un nepieciešamajā apjomā. Bioloģijā kopumā un sistēmbioloģijā uz kāda uzdevuma efektīvu realizāciju vērsta darbību kopuma aprakstu sauc par procedūru. Optimizācijas uzdevuma realizācijai nepieciešama darbību kopuma aprakstu sauc par optimizācijas procedūru (Balsa-Canto et al., 2008; Banga, 2008; Rodriguez-Acosta et al., 1999; Rodriguez-Prados et al., 2009).

Optimizācijas procedūrai vispārējā gadījumā jāatbild uz jautājumiem: kāds ir minimālais maināmo parametru skaits, ar kuriem jāmanipulē bioķīmiskā sistēmā, lai iegūtu dzīvotspējīgu organismu, kas ražo maksimāli iespējamo produkta daudzumu vai nodrošina tā labāko attiecību pret patērētās izejvielas daudzumu (Nikolaev et al., 2010; Rodriguez-Acosta et al., 1999; Rodriguez-Prados et al., 2009; Sendin et al., 2010; Vital-Lopez et al., 2006)? Kāda ir labākā darbību secība, kura varētu mūs novest līdz tādam risinājumam (Rodriguez-Acosta et al., 1999; Sendin et al., 2010)? Šādas procedūras izveide ļautu ietaupīt daudz laika un resursu, kā arī uzlabot biotehnoloģisko uzņēmumu ražību (Rodriguez-Acosta et al., 1999; Rodriguez-Prados et al., 2009). Pašlaik joprojām nav izstrādāta pietiekami detalizēta bioķīmisko tīklu vadības optimizācijas procedūra, kura ieteiktu pilnu darbību secību, sākot ar kritērija, organisma un modeļa izvēli un beidzot ar risinājuma tehniskās realizējamības pārbaudi, sadarbojoties pētniekim un ražotājiem.

### **Promocijas darba mērķis un uzdevumi**

Promocijas darba mērķis ir izstrādāt bioķīmisko tīklu datormodeļu stacionāro stāvokļu optimizācijas procedūru biotehnoloģiskās ražošanas uzlabošanai, sākot ar optimizācijas kritēriju izvirzīšanu līdz pamatošiem ieteikumiem ražošanas eksperimentu veikšanai.

Darba mērķa sasniegšanai izstrādes procesā ir jāveic vairāki uzdevumi:

- 1) veikt bioķīmisko tīklu datormodeļu stacionāro stāvokloptimizācijas tehnoloģiju un procedūru analīzi;
- 2) izveidot bioķīmisko tīklu datormodeļu stacionāro stāvokļu optimizācijas procedūru;
- 3) noskaidrot optimizācijas procedūras starpdisciplinārajā sadarbībā iesaistīto speciālistu grupas un izstrādāt koordinētas sadarbības nostādnes;

- 4) noskaidrot neperspektīvo optimizējamo parametru kombināciju atmešanas iespējas pieļaujamo risinājumu analīzes paātrināšanai;
- 5) noskaidrot faktorus, kuri ietekmē bioķīmisko tīklu modeļu stacionārā stāvokļa optimizācijas kritērija dinamiku;
- 6) izstrādāt rīku bioķīmisko tīklu modeļu stacionārā stāvokļa optimizācijas kritērija labākās vērtības dinamikas datu apstrādei un analīzei;
- 7) demonstrēt procedūras darbību, pamatojot ieteikumus rūpnieciskai etanolražošanai no glikozes, izmantojot maizes rauga (*Saccharomyces cerevisiae*) glikolīzes modeli.

### **Pētījuma metodes**

Darba izstrādē bioķīmisko procesu modelēšanā izmantotas deterministiskās kvantitatīvās dinamiskās modelēšanas metodes, pielietots SBML (*systems biology markup language*) standarts modeļa definēšanai.

Bioķīmisko modeļu optimizācijai pielietotas globālās stohastiskās optimizācijas metodes.

Optimizācijas kritērija izstrādē izmantoti svara koeficienti.

Bioķīmisko modeļu kinētisko sakarību un konstanšu pārbaudei ir pielietotas brīvpieejas datu bāzes KEGG, SABIO-RK un BRENDA.

Aprakstošās statistikas metodes izmantotas labākā risinājuma konvergences dinamikas analīzei vairāku optimizācijas eksperimentu salīdzināšanai.

Dinamisko modeļu stacionāro stāvokļu realizācijas iespējas analizētas, izmantojot modeļa dinamikas grafikus.

Maināmo parametru iespējamo kombināciju skaita novērtēšanai izmantotas kombinatorikas metodes.

### **Zinātniskais jauninājums un praktiskā vērtība**

- Izveidota bioķīmisko tīklu optimizācijas procedūra no biotehnoloģiskās ražošanas optimizācijas kritēriju izvirzīšanas līdz rūpnieciskiem testiem, īemot vērā biotehnoloģiskās ražošanas rentabilitātes aspektus.
- Izstrādātas dažādu zinātnes nozaru pārstāvju starpdisciplinārās sadarbības norādes bioķīmisko tīklu datormodeļu optimizācijas procedūrā biotehnoloģiskās ražošanas vajadzībām.
- Izstrādāta optimizācijas potenciāla atlikuma vērtēšanas metode analizējamo maināmo parametru kombināciju skaita samazināšanai gadījumos, kad atlikušo iespējamo uzlabojumu nav iemesla analizēt ekonomisku apsvērumu dēļ.

- Veikta paralēlu optimizācijas eksperimentu konvergences dinamikas statistiskā novērtējuma izstrāde (realizēta programmpunktā ConvAn).
- Izstrādāta kombinētā maināmo parametru kombināciju ranžēšanas metode, kurā tiek izmantota visu kombināciju analīze nelielam maināmo parametru skaitam kombinācijā un labākā nākamā maināmā parametra meklēšana turpmāko optimālo kombināciju atrašanai.

Darba praktisko vērtību nosaka izstrādātā bioķīmisko tīklu optimizācijas uzdevumu realizācijas procedūra. Tajā tiek norādīta darbību secība attiecībā uz bioloģiskā objekta izvēli, konkrētā objekta izpēti, t.i., optimizējamo parametru noteikšanu, kā arī cita veida informāciju par noteiktā objekta izzināšanu. Optimizācijas procedūras rezultātā ir iespējams iegūt biotehnoloģiskajā ražošanā pielietojamus organismu bioķīmisko tīklu modifikācijas ieteikumus ražošanas ekonomisko rādītāju uzlabošanai konkrētajā ražošanas procesa fāzē.

### **Darba struktūra un apjoms**

Promocijas darbs ir uzrakstīts latviešu valodā, satur anotāciju, ievadu, 4 nodaļas, secinājumus, literatūras sarakstu, 5 pielikumus, tajā skaitā 11 tabulas, 54 attēlu, 37 formulas, kopā 158 lappuses. Darbā izmantoti 171 literatūras avoti.

## **1. SISTĒMBIOLOGIJAS PIEEJA BIOPROCESU VADĪBAS PROBLĒMU RISINĀŠANĀ**

Sistēmbioloģija ir progresīva starpdisciplināra nozare, kura Latvijā attīstās saskaņā ar izvīrzītajām prioritātēm zinātnē.

Sistēmbioloģijas nozarē biologi un mediķi kopā veido aptuveni pusi no sistēmbioloģijā iesaistītajiem pētniekiem. Pārējie speciālisti pārstāv informācijas tehnoloģijas, matemātiku, sistēmanalīzi un citas ar inženierzinātnēm saistītas nozares.

Pēdējā biotehnoloģijas attīstības desmitgadē būtiskas pozīcijas ir ieņēmusi sistēmbioloģija, kuras viens no būtiskākajiem aspektiem ir informācijas tehnoloģiju izmantošana bioprocesu pētniecībā. Sistēmbioloģijas un automātiskās vadības teorijas apvienojums ir perspektīvs virziens jaunu ārstniecisko terapiju un biotehnoloģisko procesu izstrādē.

Modelēšana ir viens no priekšnoteikumiem sistēmbioloģijas pieejas realizācijai.

Bioķīmisko tīklu deterministiskie kvantitatīvie dinamiskie modeļi ir aprakstāmi ar nelineāru diferenciālvienādojumu sistēmām.

Sistēmbioloģijā ir izveidoti specializēti modeļu apmaiņas standarti (SBML, XML, MathML), kurus atbalsta vairāki specializēti programmpunkti.

## **2. BIOĶĪMISKO TĪKLU OPTIMIZĀCIJAS UZDEVUMU RISINĀŠANAS PIEEJU APSKATS**

Bioķīmisko tīklu optimizācijā izšķir divas uzdevumu klasses: parametru novērtēšanas uzdevumi (inversais uzdevums) un projektēšanas uzdevumi.

Pieaugot bioķīmisko modeļu skaitam un to lielumam, turpina attīstīties modeļu optimizācijas procedūras, kurās stohastiskās globālās optimizācijas metodes ir guvušas lielāko praktisko pielietojumu to relatīvi ātro un drošo konvergences īpašību dēļ. Šīs metodes mēdz kombinēt arī hibrīdās metodēs, kur notiek pārslēgšanās no stohastiskām globālām metodēm uz lokālajām metodēm un otrādi.

Nevienā ar bioķīmisko tīklu optimizāciju saistītā procedūrā netiek skarts dažādu nozaru pārstāvju koordinētas starpdisciplinārās sadarbības aspekts.

Bioķīmisko tīklu projektēšanas uzdevumu risināšanas gadījumā labākā risinājuma konvergences dinamika ir maz pētīta, salīdzinot ar parametru novērtēšanas uzdevumu gadījumiem.

Bioķīmisku tīklu reakciju dinamika biotehnoloģisku ražošanas uzdevumu gadījumā ir atsevišķa nelineāru diferenciālvienādojumu sistēmu uzdevumu klase, kuru iespējams aprakstīt ar ierobežotu vienādojumu tipu skaitu un kura jārisina ar stacionārā stāvokļa esamības priekšnosacījuma palīdzību.

Specializētie programmpunkti ar bioķīmisko uzdevumu risināšanai piemērotu lietotāja saskarni ļauj efektīvāk sadarboties modelētājiem, biologiem un ražotājiem, salīdzinot ar universāliem modelēšanas rīkiem.

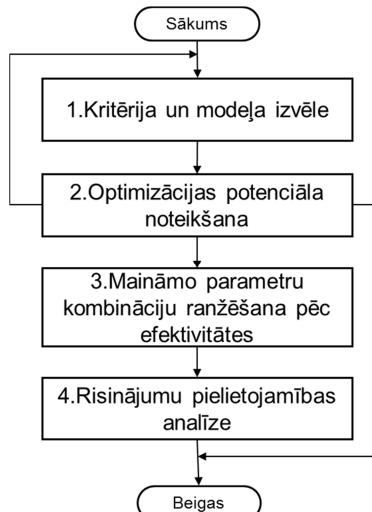
Metaboliskās vadības analīzes teorija tiek plaši pielietota piemērotāko bioprocesu projektēšanas uzdevumos labāko maināmo parametru kombināciju meklējumos.

Līdzšinējo procedūru vai metodiku starpā joprojām nav atrodama procedūra, kura aptvertu pilnu bioķīmisko tīklu projektēšanas ciklu, sācot ar kritērija izvēli un beidzot ar gatavību rūpnieciskiem testiem, analizējot visas iespējamās optimizējamo parametru kombinācijas, lai atrastu un izvērtētu procesam atbilstošākās.

### **3. OPTIMIZĀCIJAS PROCEDŪRAS DARBĪBAS UN TO SECĪBA**

Sekojot izstrādātajai bioķīmisko tīklu datormodeļu stacionāro stāvokļu optimizācijas procedūrai, ir iespējams efektīvi izmantot resursus optimizācijas laikā, jo ir iespējams izvairīties no nevajadzīgu optimizāciju uzsākšanas vai pārtraukt uzsāktu optimizācijas procesu, ja tas nav lietderīgs. Procedūras pamatprincips ir ar katru nākamo procedūras soli gūt maksimāli iespējamo informācijas daudzumu ar iespējami maz darbībām. Ar informāciju šajā gadījumā tiek saprasta informācija vai nu par perspektīviem vai neperspektīviem optimizācijas uzdevuma risinājumiem.

Ar bioķīmisko tīklu datormodeļu stacionāro stāvokļu optimizācijas procedūru var aprakstīt dažādu bioķīmisku procesu efektīvas optimizācijas darbību secību (Mozga, Stalidzans, 2011a). Līdz ar to tā ir uzskatāma par universālu attiecībā pret pētāmo organīsmu un tā modeli, jo šajā ziņā procedūrai nav ierobežojumu. Procedūra vispārīgā veidā skatāma 1. attēlā un katra posma shematiskais attēlojums un vispārīgs soļu skaidrojums ir aprakstīts turpmākajos nodaļas apakšpunktos. Procedūras detalizētājā aprakstā tiek minētas arī dažādu speciālistu grupu lomas katru soļa izpildē, lai novērstu dublēšanos, kā arī ietverta bioloģijai, ražošanai vai ekonomiskajiem apstākļiem nepiemērota vai neiespējama variantu analīze.



**1.att. Optimizācijas procedūras darbību secība**

Autora veiktā pētījuma rezultātā iesaistāmie speciālisti ir iedalīti trīs grupās – biologi, modelētāji un ražotāji. Iesaistīto speciālistu grupām nosacīti tiek piedēvētas dažādas kompetences.

**Biologi** pārzina bioķīmisko procesu mehānismus molekulārā līmenī un spēj novērtēt, pie kādiem vides parametriem ir spējīgi darboties biotehnoloģisko procesu nodrošinošie organismi un kādas izmaiņas organismu bioķīmiskajos procesos ir realizējamas ar mūsdienās pieejamajām tehnoloģijām. Biologiem jāveido priekšstati par tehnoloģiju aptuvenajām izmaksām.

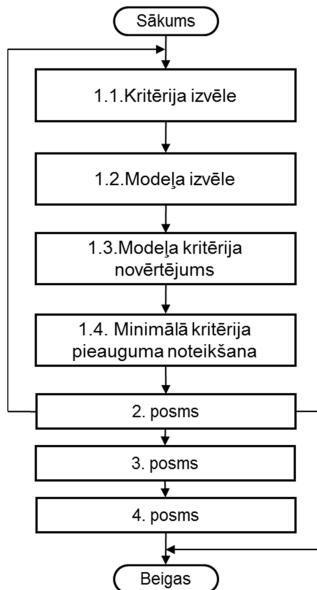
**Modelētājiem** ir pamatzināšanas darbībās ar bioķīmisko procesu modeļiem, sistēmteoriju, automātiskās vadības teoriju, kā arī labas zināšanas optimizācijas jautājumos. Modelētājiem jāorientējas pieejamo programmproduktu klāstā, lai katram uzdevumam varētu izvēlēties piemērotāko risināšanas ceļu.

**Ražotāja** termins šajā pētījumā attiecināms uz biotehnoloģiski orientēta uzņēmuma speciālistiem: 1) uzņēmuma vadītāji un ekonomisti, kuri spēj prognozēt izejvielu un produkta cenas un tirgus stāvokli, lai noteiktu optimizētā biotehnoloģiskā procesa ekonomiskos rādītājus, un 2) biotehnoloģisko iekārtu inženierus, kuri spēj novērtēt, vai optimizētais process ir realizējams uz pieejamajām iekārtām, vai arī šo nepieciešamo modifikāciju izmaksas.

Autora izveidotā optimizācijas procedūra apraksta darbības, kas sākotnēji ietver kritērija un modeļa izvēli. Protī, kurām reakcijām un ar kādiem nosacījumiem veikt optimizāciju. Tās rezultātā tiek iegūtas optimizējamo parametru modificētās skaitiskās vērtības, kuras, veicot risinājumu pielietojamības analīzi, var nodot ražotājiem ražošanas testu veikšanai. Gadījumā, ja optimizācijas darbību turpinājums varētu nest zaudējumus, tiek pārtraukta iesāktā optimizācija un piemeklēts cits bioķīmiskais process.

### **Kritērija un modeļa izvēle**

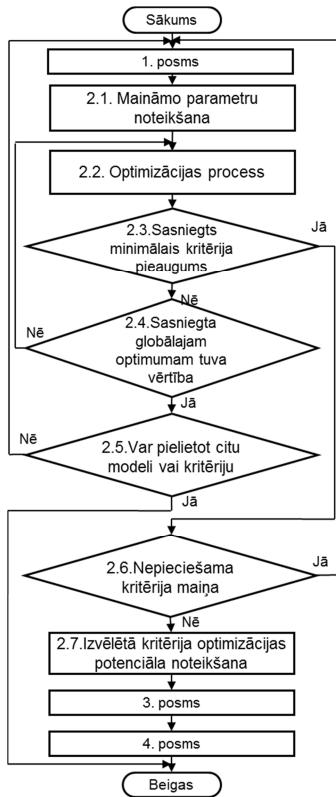
Procedūras pirms posms sākas ar kritērija un modeļa izvēli, kurš ir paredzēts procesa vispārējai izzināšanai – kāds ir biotehnoloģiskās ražošanas uzdevums, kāds ir bioķīmiskais process, kādi kritēriji ir izvirzīti. Šādu informāciju parasti sniedz ražotāji. Turpinājumā tiek skatīti modeļi, ar kādiem varētu realizēt noteiktos kritērijus, ko savukārt veic modelētājs. Šajā posmā pamatā ir ražotāja izvirzīts kritērijs atbilstoši produktam, ko vēlas ražot un, veicot ekonomiskos aprēķinus, kāds ir minimālais kritērija pieaugums, lai panāktu vēlamo biotehnoloģiskās ražošanas uzlabošanas rezultātu (sk. 2.att.).



2.att. **Kritērija un modeļa izvēle**

### **Optimizācijas potenciāla noteikšana**

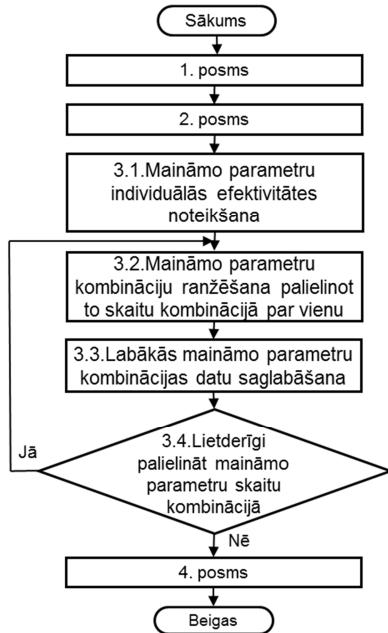
Procedūras otrs posms paredz izvēlētā modeļa uzvedības sīkāku izzināšanu (sk. 3.att.). Posma sākumā, sadarbojoties ar bioloģiem, tiek atlasīti visi modeļa parametri, kurus ir iespējams modificēt pašreizējā mikrobioloģijas tehnoloģiju attīstības stadijā. Ar attiecīgajiem parametriem tiks veiktas turpmākās optimizācijas. Tālāk tiek veikts optimizācijas process, kura laikā notiek vairākas pārbaudes, lai secinātu, vai izvēlētais modelis ar konkrētajām modificējamām reakcijām spēj sasniegt iepriekšējā posmā noteikto minimālo kritērija pieaugumu. Pozitīvas atbildes gadījumā optimizācija tiek turpināta, pretējā gadījumā jāmaina modelis vai arī kritērijs.



3.att. Optimizācijas potenciāla noteikšana

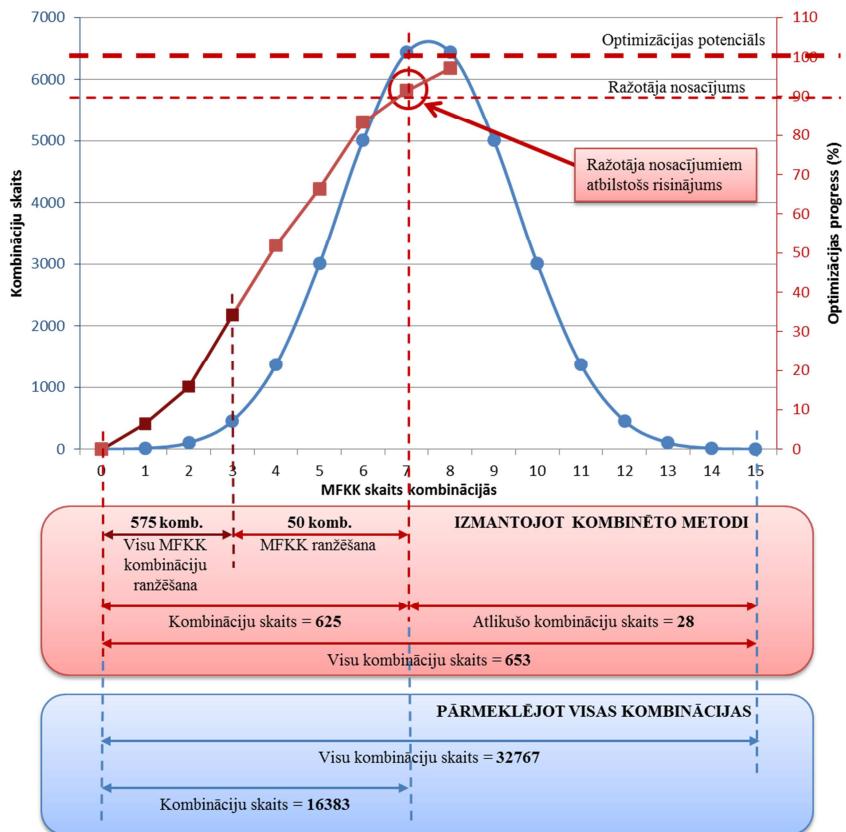
### Maināmo parametru kombināciju ranžēšana pēc efektivitātes

Procedūras trešā posma uzdevums ir efektīvāko maināmo parametru kombināciju atrāšana konkrēta maināmo parametru skaita gadījumā. Ar efektivitāti tiek saprasta maināmo parametru vai to kombināciju spēja uzlabot optimizācijas kritērija vērtību. Lielākam optimizācijas kritērija pieaugumam atbilst augstāka efektivitāte. Tieks pieņemts, ka katras reakcijas maināmo parametru koncentrācijas izmaiņu realizācija genoma līmenī izmaksā vienādu summu (procedūrā arī var tikt nemtas vērā precīzētas summas). Tajā pašā laikā ir skaidrs, ka dažādu reakciju un to kombināciju ietekme uz optimizācijas kritēriju ir ļoti atšķirīga. Tādēļ jāatrod maināmo parametru kombinācija ar iespējamī mazāko maināmo parametru skaitu, kas spētu izmantot lielāko daļu no pētāmā procesa kopējā potenciāla, kurš jau ir noteikts solī „2.7. Izvēlētā kritērija optimizācijas potenciāla noteikšana” (sk. 4.att.).



4.att. **Maināmo parametru kombināciju ranžēšana pēc efektivitātes**

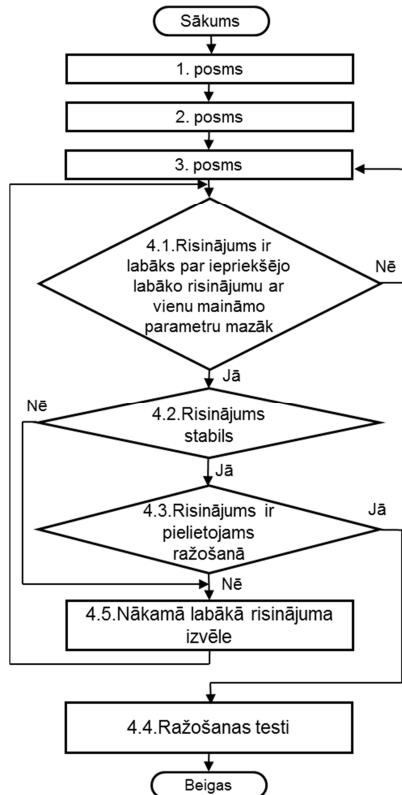
5. attēlā ir novērtējams arī optimizācijas potenciāla atlikuma vērtēšanas metodes pielietošanas rezultāts: optimizācija tika apturēta pie septiņiem MFKK kombinācijā, jo optimizācijas potenciāla atlikums nebija pietiekošs ekonomiskā labuma gūšanai biotehnoloģiskajā ražošanas procesā. Rezultātā konstatēts, ka nav vērts analizēt 16384 kombinācijas (pilnās pārmeklēšanas gadījums), tādējādi būtiski samazinot veicamā darba un patēriņamā laika apjomu.



5.att. Optimizācijas potenciāla atlikuma vērtēšanas metodes pielietojums analizējamo kombināciju skaita samazināšanai

### Risinājumu pielietojamības analīze

Procedūras ceturtajā posmā veicama atrasto risinājumu pielietojamības analīze un iegūto optimizācijas procesu gala parametru rezultātu eksperimentāla testēšana ražošanā (sk. 6.att.). Ceturta posma ietvaros tiek pārbaudīts, vai aistrastais risinājums ne tikai uzrāda stacionāra stāvokļa esamību, bet ir arī pietiekoši stabils realizācijai ar ražošanas iekārtām par pieņemamām izmaksām. Papildus nepieciešams pārbaudīt, vai stacionārā stāvokļa parametri (vielu koncentrācijas, temperatūra,  $pH$  vērtības utt.) ir realizējami.



6.att. Atrasto risinājumu pielietojamības analīze

#### 4. KRITĒRIJA VĒRTĪBAS KONVERGENCES ANALĪZE STACIONĀRĀ STĀVOKĻA OPTIMIZĀCIJAS EKSPERIMENTA LAIKĀ

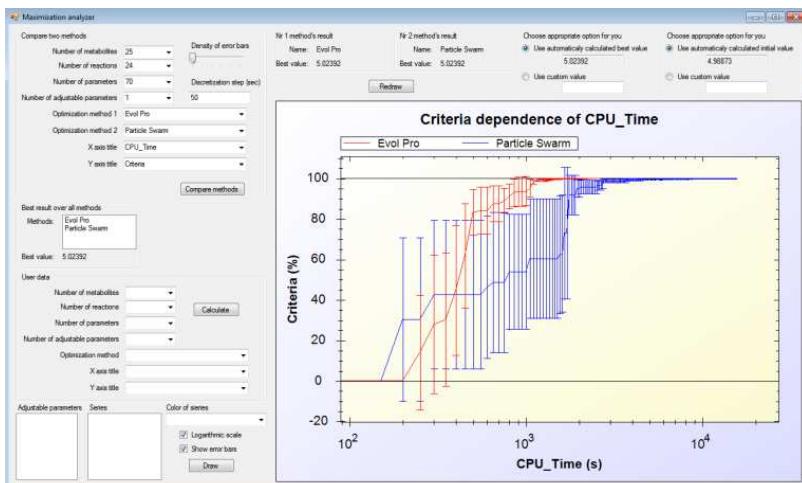
Skaitliskā optimizācija pielietojuma piemēra risināšanas gaitā parādās kā laikieltpīgākais process ar neviennozīmīgu rezultātu tā apstākļa dēļ, ka skaitliskās metodes tikai konverģē uz globālo optimumu, kura vērtība nav zināma. Optimizācijas eksperimentu ilgums var svārstīties robežās no dažām minūtēm līdz pat 60-90 stundām dažu desmitu reakciju lielam modelim (Nikolaev, 2010). Praktiski tas izpaužas tādējādi, ka, veicot optimizāciju, nav skaidrs, kurā brīdī optimizācijas procesu varētu beigt, uzskatot, ka būtisks kritērija vērtības uzlabojums nav gaidāms. Tā kā ir pieejamas daudzas optimizācijas metodes un katrai no tām mēdz būt parametri, kuru maiņa ietekmē to efektivitāti, metožu pielietojamību, to konvergences ātrums un stabilitāte ir ļoti svarīgs optimizācijas procedūras efektivitāti ietekmējošs faktors.

## Konvergences analīzes rīks ConvAn

Balstoties uz promocijas darba nodaļās aprakstītajiem eksperimentālajiem datiem, var secināt, ka optimizācijas kritērija konvergēnci ir grūti prognozējama un ir nepieciešami tās statistiskās analīzes rīki. Darba ietveros tika izstrādāta rīka koncepcija un izvirzītas funkcionalitātes prasības. Sadarbībā ar Andreju Kostrominu tika izstrādāts programmatūras rīks ConvAn (Kostromins et al., 2012) ar sekojošu funkcionalitāti:

- optimizācijas datu glabāšana, rediģēšana, dzēšana;
- optimizācijas metožu reģistrēšana;
- uz optimizāciju nosacījumiem balstītu profilu veidošana;
- optimizācijas metožu salīdzināšana pie vienādiem nosacījumiem;
- konkrētas metodes analizēšana pie dažādiem nosacījumiem.

Konvergences analīzes rīka ConvAn nosaukums veidots no angļu valodas vārdiem *Convergence analysis*, kas tulkojumā nozīmē konvergences analīze. ConvAn rīks ir brīvpieejas programmatūra, kuras jaunākās versijas, paraugdati, instrukcijas, video pamācība un dokumentācija atrodama interneta vietnē <http://www.biosystems.lv/convan>. ConvAn rīks darbojas Windows un Linux operētāsistēmās (sk. 7.att.).



7.att. Rīka ConvAn metožu salīdzināšana maksimizācijas uzdevumā

Konvergences analīzes rīka (ConvAn) ieejas parametri ir divu vektoru matrica, kurā viens vektors ir optimizācijas laiks un otrs vektors ir labākā risinājuma kritērija vērtība attiecīgajā laika momentā. Kā papildu dati tiek ievadīti dati par modeli (virsotņu skaits, reakciju skaits, nodalījumu skaits), kā arī optimizācijas

uzdevuma parametri (maināmo parametru skaits, maināmo parametru diapazons/iespējamo vērtību skaits).

ConvAn izejas parametri ir dati par konkrētā modeļa konvergences ātruma statistisko novērtējumu pielietotajai optimizācijas metodei. Šos datus var izmantot citu optimizācijas eksperimentu konvergences dinamikas analīzei. Var tikt vērtēta arī optimizācijas metodes un tās parametru piemērotība konkrētā uzdevuma risināšanai.

## SECINĀJUMI

### Galvenie darba rezultāti

Ir izstrādāta bioķīmisko tīklu datormodeļu stacionāro stāvokļu optimizācijas procedūra biotehnoloģiskās ražošanas uzlabošanai, sākot ar optimizācijas kritēriju izvirzīšanu līdz pamatotiem ieteikumiem ražošanas eksperimentu veikšanai.

1. Ir veikta bioķīmisko tīklu datormodeļu optimizācijas tehnoloģiju un procedūru analīze.

Citu autoru piedāvātās bioķīmisko tīklu datormodeļu optimizācijas procedūras katrā ietver tikai daļu no nepieciešamajām darbībām, lai nonāktu no biotehnoloģiskās ražošanas uzlabošanas līmena uzdevuma nostādnes līdz labākā risinājuma ieteikumam.

Nav atrasta bioķīmisko tīklu datormodeļu optimizācijas procedūra biotehnoloģiskās ražošanas vajadzībām, kurā nemta vērā koordinēta starpdisciplināra sadarbība, kas ļautu nemt vērā dažādu speciālistu grupu savlaicīgi izteiktus ierosinājumus vai ierobežojumus, kā arī atmest kādai no speciālistu grupām nepieņemamus vai neperspektīvus risinājumus.

2. Izveidota bioķīmisko tīklu datormodeļu stacionāro stāvokļu optimizācijas procedūra.

Izstrādātā bioķīmisko tīklu datormodeļu stacionāro stāvokļu optimizācijas procedūra sastāv no četriem secīgi izpildāmiem posmiem: kritērija un modeļa izvēle, optimizācijas potenciāla noteikšana, maināmo parametru kombināciju ranžēšana pēc efektivitātes un risinājumu pielietojamības analīze. Darbību secība izvēlēta pēc principa: katrs nākamais procedūras solis ir labākais pēc kritērija maksimāls jauniegūtās informācijas daudzums, veicot minimālu darbību apjomu.

3. Noskaidrotas optimizācijas procedūras starpdisciplinārajā sadarbībā iesaistīto speciālistu grupas un izstrādātas koordinētas sadarbības nostādnes.

Optimizācijas procedūrā iesaistīti trīs grupu speciālisti: biologi, modelētāji un ražotāji (dažādi ražošanas uzņēmuma pārstāvji). Izveidotajā bioķīmisko tīklu optimizācijas procedūrā ir iestrādātas norādes par dažādu procedūras izpildes gaitā nepieciešamo lēmumu pieņemšanā iesaistāmajām speciālistu grupām un to uzdevumiem.

4. Noskaidrotas neperspektīvo optimizējamo parametru kombināciju atmešanas iespējas pieļaujamo risinājumu analīzes paātrināšanai.

Izstrādāta optimizācijas potenciāla atluka vērtēšanas metode analizējamo maināmo parametru neperspektīvo kombināciju atmešanai gadījumos, kad atlkušo iespējamo uzlabojumu nav iemesla analizēt ekonomisku apsvērumu dēļ.

5. Noskaidroti faktori, kuri ietekmē bioķīmisko tīklu modeļu stacionārā stāvokļa optimizācijas kritērija dinamiku.

Bioķīmisko tīklu modeļu stacionārā stāvokļa optimizācijas uzdevumu risināšanā lielāko popularitāti ieguvuši globālie stohastiskie skaitliskās optimizācijas algoritmi, kuri ir universāli pielietojami neatkarīgi no modeļa īpatnībām un nelinearitātes veida.

Stohastisko globālo optimizācijas metožu konvergences ātrums un globālā optimuma sasniegšana ir atkarīga gan no maināmo parametru skaita, gan no izvēlētajiem maināmajiem parametriem konstanta maināmo parametru skaita gadījumā, gan arī no izvēlētās optimizācijas metodes.

6. Izstrādāts rīks bioķīmisko tīklu modeļu stacionārā stāvokļa optimizācijas kritērija labākās vērtības dinamikas datu apstrādei un analīzei.

Ir izveidots rīks ConvAn optimizācijas konvergences ātruma izvērtēšanai un prognozēšanai. Vairākas optimizācijas ar identiskiem sākuma nosacījumiem tiek izvērtētas ar statistiskām metodēm. Uz analīzes pamata var prognozēt konvergences ātrumu līdzvērtīgam modelim, izmantojot to pašu metodi. ConvAn var izmantot optimizācijas kritērija konvergences ātruma novērtēšanai un prognozei gan parametru novērtēšanas, gan projektēšanas uzdevumu gadījumos.

7. Demonstrēta procedūras darbība, pamatojot ieteikumus rūpnieciskai etanola ražošanai no glikozes, izmantojot maizes rauga (*Saccharomyces cerevisiae*) glikolīzes modeli.

Optimizācijas procedūras izpildes rezultātā Hinnes un līdzautoru (Hynne et al., 2001) izveidotajam maizes rauga (*Saccharomyces cerevisiae*) glikolīzes modelim ir noteiktī septiņu reakciju fermentu koncentrāciju maiņas koeficienti, pieļaujot izmaiņas diapazonā 0,01 – 11,0. Manipulējamās reakcijas un to koeficienti (iekavās) ir: glikozes patēriņš (11,0), heksokināze (2,73), fosfofruktokināze (2,02), piruvātdekarboksilāze (5,93),

alkoholdehidrogenāze (11,0), rezerves vielas (0,01), ATF patēriņš (11,0). Modelis prognozē, ka izmaiņu ieviešanas rezultātā etanola produktivitāte palielinās no 0,80 uz 2,05 mmol/min, bet glikozes patēriņš palielinās no 0,83 uz 1,12 mmol/min, un samazinās blakusproduktu rašanās glicerīnam no 0,085 mmol/min uz 0,028 mmol/min, acetaldehīdam no 0,065 mmol/min uz 0,021 mmol/min un cianīdam no 0,020 mmol/min uz 0,007 mmol/min. Rezultātā tiek uzlabots gan produkta iznākums no izejvielas vienības no 0,97 uz 1,83, gan arī 2,6 reizes palielinās etanola produktivitāte.

### **Secinājumi un attīstības perspektīvas**

Pieaugot bioķīmisko modeļu skaitam un to lielumam, turpina attīstīties modeļu optimizācijas procedūras, kurās stohastiskās globālās optimizācijas metodes ir guvušas lielāko praktisko pielietojumu to relatīvi ātro un drošo konvergences īpašību dēļ. Šīs metodes mēdz kombinēt arī hibrīdās metodēs, kur notiek pārslēgšanās no stohastiskām globālām metodēm uz lokālajām metodēm un otrādi.

Bioķīmisku tīklu reakciju dinamika biotehnoloģisku ražošanas uzdevumu gadījumā ir atsevišķa nelineāru diferenciālvienādojumu sistēmu uzdevumu klase, kuru iespējams aprakstīt ar ierobežotu vienādojumu tipu skaitu un kura jārisina ar stacionārā stāvokļa esamības priekšnosacījuma palīdzību.

Piemērota optimizācijas kritērija izvēle projektēšanas uzdevuma gadījumā ir veicama iteratīvi, sadarbojoties biologiemi, modelētājiem un ražotājiem, un ir atkarīga ne tikai no ekonomiskiem un tehnoloģiskiem faktoriem, bet arī no modeļa apjoma un tā īpatnībām.

Maināmo modeļa parametru ranžēšana bioķīmisko tīklu modeļu stacionāro stāvokļu optimizācijas gadījumā atpaliek pēc efektivitātes no maināmo modeļa parametru kombināciju ranžēšanas.

Izstrādātā kombinētā ranžēšanas metode lielam maināmo parametru skaitam piedāvā izmantot visu kombināciju analīzi, ja parametru skaits kombinācijā ir mazs, un pāriet uz parametru ranžēšanu, pārmeklējot labāko nākamajā kombinācijā iesaistāmo parametru no atlikušajiem parametriem. Šāda pieeja nevar garantēt labāko kombināciju atrašanu lielākam parametru skaitam, taču rezultāts ir sasniedzams saprātīgā laikā. Gadījumā, ja visu kombināciju skaits ir neliels vai, ja pietiek resursu visu iespējamo kombināciju pārbaudei, jāpārbauda visas kombinācijas.

Nemainīga modeļa, sākotnējo nosacījumu un optimizācijas metožu iestatījumu apstākļos stohastisko globālo optimizācijas metožu konvergences ātrums un

globālajam optimumam tuvas vērtības sasniegšana var būt atkarīga no maināmo parametru skaita, no to īpašībām, kā arī no izvēlētās optimizācijas metodes.

Kopējā optimizācijas potenciāla gadījumā, pielietojot optimizācijas procedūru un mainot 15 fermentu koncentrācijas, ir izdevies palielināt etanola produktivitāti no 0,80 mmol/min uz 2,27 mmol/min un uzlabot produkta iznākumu no izejvielas vienības no 0,97 uz 1,97, un samazināt blakusprodukta rašanos glicerīnam no 0,085 mmol/min uz 0,007 mmol/min, acetaldehīdam no 0,065 mmol/min uz 0,005 mmol/min un cianīdam no 0,020 mmol/min uz 0,002 mmol/min.

**Kā attīstības perspektīvas** var iezīmēt vairākus darbības virzienus.

1. Izstrādāt kritērijus, pēc kuriem varētu veikt optimizācijas eksperimenta beigšanu, izmantojot globālās stohastiskās metodes, balstoties uz vairāku paralēlu optimizāciju konvergences dinamikas statistisko analīzi.
2. Maināmo parametru kombināciju analīzei nepieciešams izveidot programmproduktu, kas automātiski veiktu visu iespējamo kombināciju pārbaudi un to ranžēšanu pēc optimizācijas kritērija dažādam maināmo parametru skaitam, izmantojot optimizācijas eksperimenta beigšanas kritērijus un brāķējot neperspektīvos risinājumus pēc to vidējām konvergences dinamikas līknēm.
3. Meklēt iespējas izstrādātos enzīmu koncentrāciju maiņas ieteikumus pārbaudīt bioloģiskos eksperimentos.
4. Salīdzināt brīvpieejas programmpunktos pieejamās optimizācijas metodes pēc to labākā risinājuma konvergences dinamikas, izmantojot ConvAn rīku.

## PARTICULARS

**Executed at:** Latvia University of Agriculture, Faculty of Information Technologies, Computer Systems Department.

**Scientific Adviser of the Doctoral Thesis:** Dr.sc.ing. Egils Stalidzāns, Associate Professor, Latvia University of Agriculture.

**Consultant of the Doctoral Thesis:** Dr.habil.biol. Pēteris Zikmanis, Senior Researcher, Institute of Microbiology and Biotechnology, University of Latvia.

### Scientific approbation of the thesis at the final stage.

1. Approbated at the seminar on 11 June, 2009.
2. Discussed and approbated at ESF project „Establishment of Latvian interdisciplinary interuniversity scientific group of systems biology” (agreement Nr. 2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128)
3. Accepted as fully complete at the Promotion Council of Faculty of Information Technologies, Latvia University of Agriculture on 29 February, 2012.

**The thesis was approved** at the expanded academic session of the Computer Systems Department of the Faculty of Information Technologies of the Latvia University of Agriculture on 25 May, 2011. Minutes No 1.

The doctoral thesis was produced with the assistance of the **European Social Fund (ESF)** project „Atbalsts doktorantūras studijām un pēcdoktorantūras pētījumiem inženierzinātnēs un mežzinātnē” (agreement No 2004/0004/VPD1/ESF/APIA/04/NP/3.2.3.1/0001/0005/0067) and **ESF** project „Establishment of Latvian interdisciplinary interuniversity scientific group of systems biology” (agreement No 2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128).

### Official Reviewers:

1. Professor of the Latvia University Dr.sc.comp. Juris Vīksna;
2. Professor of the Riga Technical University Dr.habil.sc.ing. Jurijs Merkurjevs;
3. Professor of the Abo Akademi (Turku, Finland) PhD Ion Petre.

**The defence of the doctoral thesis** will take place at the open session of the Promotion Council in the field of Information Technologies of LUA at 1 p.m. on 22 May 2012, Room No.218, Faculty of Information Technologies Lielā 2, Jelgava.

The thesis can be accessed at the LUA Fundamental Library, Lielā 2, Jelgava, and at <http://llufb.llu.lv/llu-theses.htm>.

**You are welcome to send your comments**, signed and in a scanned form to secretary of Promotion Council - Lielā 2, Jelgava, LV-3001; phone: (+371) 63005621; e-mail: tatjana.tabunova@llu.lv.

**Council Secretary:** lector, Mg.paed. Tatjana Tabunova.

## TABLE OF CONTENTS

PARTICULARS .....	24
APPROBATION OF PhD THESIS .....	26
INTRODUCTION.....	29
Theme topicality .....	29
The aim and tasks of PhD thesis .....	30
Research methods .....	31
Scientific novelty and practical value .....	31
PhD thesis structure and volume.....	32
1. SYSTEMS BIOLOGY APPROACH IN DEALING WITH BIOPROCESS	
CONTROL PROBLEMS .....	32
2. REVIEW OF APPROACHES USED IN SOLVING BIOCHEMICAL	
NETWORK OPTIMIZATION TASKS.....	33
3. OPTIMIZATION PROCEDURE ACTIONS AND THEIR SEQUENCE .....	34
Choice of criterion and model.....	36
Determination of the optimization potential .....	36
Ranking variable parameter combinations by efficiency .....	37
Analyzing the application of solutions.....	39
4. ANALYSIS OF THE CRITERION VALUE CONVERGENCE DURING	
STEADY STATE OPTIMIZATION EXPERIMENT .....	40
Convergence analysis tool ConvAn.....	41
CONCLUSIONS.....	42
The main PhD thesis results.....	42
Conclusions and development prospects .....	44
BIBLIOGRAPHY .....	46

## APPROBATION OF PhD THESIS

The research results of PhD thesis are presented in the following publications:

1. Mozga I., Stalidzāns E. (2006) Acquisition of dynamic parameters of biological process from experimental data. International scientific conference “Information Technologies for Rural Development”, pp. 131-140. (Indexed in the AGRIS database).
2. Stalidzāns E., Gailums A., Mozga I., (2006) Life cycle of artificial and biological control system. International conference. International scientific conference “Information Technologies for Rural Development”, pp. 164-172. (Indexed in the AGRIS database).
3. Grunde-Zeiferts U., Mozga I., Žukova T., Stalidzāns E. (2006) Therapy modelling combining methods of systems biology and automatic control theory. International scientific conference “Animals. Health. Food hygiene”, pp. 70-74. (Indexed in the AGRIS database).
4. Stalidzāns E., Mozga I., (2006) Control oriented dynamic modelling of biological object. Lithuanian University of Agriculture, Biometrics and information technologies in agriculture: research and development, pp. 11-14.
5. Mozga I., Stalidzāns E. (2007) Application of dynamic models of glycolysis developing control system. 6th International Scientific Conference „Engineering for Rural Development”. pp. 119-123. (Indexed in the AGRIS database).
6. Mozga I., Grunde-Zeiferts U., Sudars S., Stalidzans E. (2007) Life Cycles and Competition in Modelling of Artificial and Biological Control System. The European Simulation and Modelling Conference, pp. 533-536. (Indexed in the AGRIS, EBSCO, Thomson Reuters database).
7. Lagzdina L., Mozga I., Stalidzans E. (2008) Algorithm for bioethanol yield optimisation in *Saccharomyces cerevisiae* glycolysis using ‘COPASI. International scientific conference “Applied Information and Communication Technologies”, pp. 95-101. (Indexed in the AGRIS, EBSCO database).
8. Mozga I., Kostromins A., Stalidzans E. (2011) Forecast of numerical optimization progress of biochemical networks. 10th International Scientific Conference „Engineering for rural development”, pp. 103-108. (Indexed in the AGRIS, EBSCO, Scopus, Thomson Reuters database).
9. Mozga I., Stalidzans E. (2011a) Optimization protocol of biochemical networks for effective collaboration between industry representatives, biologists and modellers. 9th Annual Industrial Simulation Conference “ISC’2011”, pp. 91-96. (Indexed in the Thomson Reuters database).
10. Mozga I., Stalidzans E. (2011b) Convergence Dynamics of Biochemical Pathway Steady State Stochastic Global Optimization, 12th IEEE

- International Symposium on Computational Intelligence and Informatics, CINTI 2011 - Proceedings, pp. 231-235. (Indexed in the IEEE Xplore, Scopus database).
11. Mozga I., Stalidzans E. (2011c) Convergence Dynamics of Biochemical Models To The Global Optimum, 3rd IEEE International Conference on e-Health and Bioengineering, EHB 2011 – Proceedings, pp. 227-230. (Indexed in the IEEE Xplore, Scopus database).
  12. Kostromins A., Mozga I., Stalidzans E. (2012) ConvAn: a convergence analyzing tool for optimization of biochemical networks. Biosystems, 108(1-3), pp. 73-77. Elsevier. doi:10.1016/j.biosystems.2011.12.004 (Indexed in the IEEE Xplore, Scopus database).

The research results were presented at the following conferences:

1. Mozga I., Stalidzāns E. Acquisition of dynamic parameters of biological process from experimental data. Jelgava, Latvia. 19-20 October 2006.
2. Stalidzāns E., Gailums A., Mozga I. Life cycle of artificial and biological control system. Jelgava, Latvia. 19-20 October 2006.
3. Grunde-Zeiferts U., Mozga I., Žukova T., Stalidzāns E. Therapy modelling combining methods of systems biology and automatic control theory. Jelgava, Latvia. 10 November 2006.
4. Stalidzāns E., Mozga I. Control oriented dynamic modelling of biological object. Kaunas, Lithuania. 24-25 November 2006.
5. Mozga I., Acquisition methods of SBML biomodels elements on the Internet. Jelgava, Latvia. 16-18 May 2007.
6. Mozga I., Stalidzāns E. Application of dynamic models of glycolysis developing control system. Jelgava, Latvia. 24-25 May 2007.
7. Mozga I., Grunde-Zeiferts U., Stalidzāns E. Glycolysis control modeling using Systems Biology Markup Language (SBML) software. Parnumaa, Estonia. 26-30 August 2007.
8. Mozga I., Stalidzans E., Metabolic Control Analysis tools in SBML modelling software. NJF seminar „Modelling in Agriculture”, Jelgava, Latvia, 18-20 October 2007.
9. Mozga I., Grunde-Zeiferts U., Sudars S., Stalidzans E. Life Cycles and Competition in Modelling of Artificial and Biological Control System. The European Simulation and Modelling Conference 2007, St.Julian's, Malta, 22-24 October 2007.
10. Lagzdina L., Mozga I., Stalidzans E., Algorithm for bioethanol yield optimisation in *Saccharomyces cerevisiae* glycolysis using ‘COPASI’. International scientific conference “Applied Information and Communication Technologies”, Jelgava, Latvia, 10-12 April 2008.
11. Brusbardis V., Stalidzans E., Mozga I. Biological and artificial control systems as competitors. Summer school on systems biology for medical applications, held in Tenerife, Spain, 30 September - 2 October 2008.

12. Lagzdina L., Mozga I., Stalidzans E., Algorithm for bioethanol yield optimisation in *Saccharomyces cerevisiae* Glycolysis using Copasi. Summer school on systems biology for medical applications, held in Tenerife, Spain, 30 September - 2 October 2008.
13. Mozga I., Stalidzans E., Computational optimisation of enzyme concentrations. Systems Biochemistry – Linked Focused Meeting, University of York, York, UK, 22-24 March 2010.
14. Mozga I., Kostromins A., Stalidzans E., Forecast of numerical optimization progress of biochemical networks. 10th International Scientific Conference „Engineering for rural development”, Jelgava, Latvia, 26-27 May 2011.
15. Mozga I., Stalidzans E., Optimization protocol of biochemical networks for effective collaboration between industry representatives, biologists and modellers. The Industrial Simulation Conference 2011, Venice, Italy, 6-8 June 2011.
16. Mozga I., Stalidzans E. Convergence Dynamics of Biochemical Pathway Steady State Stochastic Global Optimization. 12th IEEE International Symposium on Computational Intelligence and Informatics, CINTI 2011, Budapest, Hungary, 21-22 November 2011.
17. Mozga I., Stalidzans E. Convergence Dynamics of Biochemical Models To The Global Optimum, 3rd IEEE International Conference on e-Health and Bioengineering, EHB2011, Iasi, Romania, 24-26 November 2011.

The research results were presented at the following seminars:

1. Mozga I. General characteristics of the glycolysis process, Seminar of Biosystems group, Jelgava, Latvia, 9 October 2006.
2. Mozga I. COPASI software simulations of glycolysis process, Seminar of Biosystems group, Jelgava, Latvia, 30 October 2006.
3. Mozga I. Comparison of different models of glycolysis, Seminar of Biosystems group, Jelgava, Latvia, 6 November 2006.
4. Mozga I. Comparative analysis of glycolysis models, Seminar of Biosystems group, Jelgava, Latvia, 13 November 2006.
5. Mozga I. The forecast of optimization rate and accuracy of biochemical networks, Seminar of Biosystems group, Jelgava, Latvia, 12 February 2009.
6. Mozga I. Biochemical network modeling with CellDesigner, Seminar of Biosystems group, Jelgava, Latvia, 21 October 2009.
7. Mozga I. Systems biology tool SBtoolbox2 in Matlab software, Seminar of Biosystems group, Jelgava, Latvia, 3 March 2010.
8. Mozga I. Determination of optimization criteria in COPASI software, Seminar of Biosystems group, Jelgava, Latvia, 3 November 2010.

9. Mozga I. Optimization of biochemical networks, determination of criteria, time and accuracy forecast tool, Seminar of Biosystems group, Jelgava, Latvia, 1 December 2010.

Participation in the projects related to PhD thesis development:

1. Project of Latvian Council of Science Nr.08.2178 „Rationalization of biofuel production using dynamic models of biochemical reaction networks” (in 2008, researcher)
2. „Establishment of Latvian interdisciplinary interuniversity scientific group of systems biology”  
Nr.2009/0207/1DP/1.1.1.2.0/09/APIA/VIAA/128, (from 2009, researcher)

Titles and awards associated with the results of PhD thesis:

The Young Researcher, 3rd IEEE International Conference on e-Health and Bioengineering, EHB 2011, Iasi, Romania, 24-26 November 2011.

## **INTRODUCTION**

### **Theme topicality**

Biotechnology is a use of any living organism (or biological systems in general) to promote human welfare. Biotechnology is used in food production, waste degradation, medicine, and many other areas (Reiss, 2005).

In the European Union (EU) there is a continuously increasing interest concerning the biofuel implementation in Europe, mainly because of environmental protection and energy supply security reasons (Kondili, Kaldellis, 2007). In this context, the EU strongly encourages the use of biofuels through a number of Directives (EC 2003).

Information technologies have become one of the preconditions for the development of biotechnology; their application is determined by several reasons: storage and processing of large amounts of data, similarity analysis, data mining tasks and a variety of modelling applications. Modelling task is primarily to gather knowledge and convert them into such form it could be compared with experimental data, confirming or excluding any prior assumptions (Mauch et al., 2001). Models are also used for forecasting the impact of various changes in a system on the modelled process. This approach is widely used in metabolic engineering to search for changes in the system that are needed to improve its characteristics (Keasling, 2010; Sendin et al., 2010).

Cellular biochemical network optimization is a complex and time consuming task, especially with an increase in model size and number of variable parameters. However, in a case of cellular biochemical networks a number of regularities are valid, the use of which can simplify the task to be solved. The space of task solutions can also be limited by the information available in databases on different types of experiments. Therefore, in the course of optimization a set of varied measures must be carried out, during which the numerical optimization methods are just one of the instruments to be applied at the right time and amount needed. Both in biology in general and systems biology a description of a set of actions aimed at effective implementation of a task is called the procedure. A description of a set of actions required for implementation of optimization task is called the optimization procedure (Balsa-Canto et al., 2008; Banga, 2008; Rodriguez-Acosta et al., 1999; Rodriguez-Prados et al., 2009).

Optimization procedure in general case must answer the following questions: what is the minimum number of variable parameters, which should be manipulated with in biochemical system in order to obtain a viable organism that produces the maximum possible quantity of the product or provides the best ratio of product vs. quantity of consumed substrate (Nikolaev et al., 2010; Rodriguez-Acosta et al., 1999; Rodriguez-Prados et al., 2009; Sendin et al., 2010; Vital-Lopez et al., 2006)? What is the best sequence of actions that could lead us to this solution (Rodriguez-Acosta et al., 1999; Sendin et al., 2010)? Creation of such procedure would save a lot of time and resources, as well as improve the productivity of biotechnology companies (Rodriguez-Acosta et al., 1999; Rodriguez-Prados et al., 2009). At present, the steady state optimization procedure of biochemical networks, which would describe a full sequence of actions from the choice of criterion, organism, and model, through to testing of solution's technical feasibility in collaboration with scientists and industrialists from different industries, is yet to be developed.

### **The aim and tasks of PhD thesis**

The aim of the PhD thesis is to develop steady state optimization procedure of biochemical network computer models for improvement of biotechnological production starting from the setting of optimization criteria up to justified recommendations for production experiments.

In order to achieve the aim of the PhD thesis several tasks must be completed:

- 1) analyze optimization technology and procedure of biochemical network computer models,
- 2) develop steady state optimization procedure of biochemical network computer models,

- 3) determine groups of experts involved in the optimization procedure interdisciplinary collaboration and develop integrated collaboration references,
- 4) determine possibilities for eliminating non-perspective optimization parameter combinations in order to accelerate analysis of admissible solutions,
- 5) determine factors influencing dynamics of steady state optimization criterion of the biochemical network models,
- 6) develop a tool for dynamics data processing and analysis of the best steady state optimization criterion value of the biochemical network models,
- 7) demonstrate operation of procedure by recommendations for industrial production of ethanol from glucose using the baker's yeast (*Saccharomyces cerevisiae*) glycolysis model.

### **Research methods**

During development of PhD thesis deterministic quantitative dynamic modelling methods were used in modelling biochemical processes, and SBML (Systems Biology Markup Language) standard was applied to model definition.

Global stochastic optimization methods were used for optimization of biochemical models.

Weight coefficients are used in developing optimization criterion.

Open-access databases KEGG, SABIO-RK and BRENDA are applied in verification of kinetic relations and constants of biochemical models.

Descriptive statistical methods are used in analyzing the best solution of convergence dynamics for comparison of several optimization experiments.

Implementation options for the steady state dynamic models are analyzed using graphs of model dynamics.

Combinatorial methods are used in evaluating the number of possible combinations of variable parameters.

### **Scientific novelty and practical value**

- Biochemical network optimization procedure was developed from the formulation of biotechnological production optimization criteria up to the industrial tests taking into account the profitability aspects of biotechnological production.

- Interdisciplinary collaboration references were developed by representatives of various scientific disciplines in the optimization procedure of biochemical network computer models for biotechnological production needs.
- Method for evaluating the remaining range of the optimization potential was developed for reducing a number of combinations of variable parameters to be analyzed in cases when there is no reason for analyzing the remaining potential improvement due to economic reasons.
- Statistical evaluation of convergence dynamics of parallel optimization experiments was developed (implemented in software product ConvAn).
- Combined ranking method of variable parameter combinations was developed, where both the analysis of all combinations for a small number of variable parameters in a combination and the searching of the next best variable parameter to find the future optimum combinations is used.

The developed procedure for the implementation of biochemical network optimization tasks determines the practical value of PhD thesis. It shows the sequence of actions regarding choice of a biological object, research of the specific object, i.e., setting of optimization parameters, as well as other information on the specified object. In the result of the optimization procedure, recommendations applicable to biotechnological production for modifying biochemical networks of organisms can be obtained to improve industrial economic indicators in the specific phase of production process.

### **PhD thesis structure and volume**

The PhD thesis is written in Latvian containing abstract, introduction, 4 chapters, conclusions, bibliography, and 5 annexes, including 11 tables, 54 pictures, and 37 formulae, 158 pages in total. 171 literature sources were used.

## **1. SYSTEMS BIOLOGY APPROACH IN DEALING WITH BIOPROCESS CONTROL PROBLEMS**

Systems biology is a progressive interdisciplinary field which in Latvia develops under priorities set in science.

In systems biology field biologists and physicians together form approximately one half of the researchers involved in systems biology. Other experts represent information technologies, mathematics, system analysis, and other engineering-related fields.

During the last decade of biotechnology development a significant position is taken by systems biology, the key aspect of which is the use of information technologies in bioprocess research. The combination of systems biology and automatic control theory is a promising direction to the development of new medical therapies and biotechnological processes.

Modelling is one of the preconditions for implementing systems biology approach.

Deterministic quantitative dynamic models of biochemical networks are described by nonlinear differential equation systems.

In systems biology there are specialized model exchange standards (SBML XML, MathML) supported by a number of specialized software products.

## **2. REVIEW OF APPROACHES USED IN SOLVING BIOCHEMICAL NETWORK OPTIMIZATION TASKS**

In optimization of biochemical networks two classes of tasks are distinguished: parameter evaluation tasks (inverse task) and design tasks.

While the number of biochemical models and their size increases, the model optimization procedures continue to develop, where mostly the stochastic global optimization methods are applied due to their relatively quick and safe convergence properties. These methods may be combined also in hybrid methods, where there is a shift from stochastic global methods to local methods and vice versa.

None of the procedures related to biochemical network optimization addresses the aspect of coordinated interdisciplinary collaboration among representatives of different fields.

In the case of solving design tasks of biochemical networks the best solution convergence dynamics is not studied a lot in comparison with the cases of parameter evaluation tasks.

The dynamics of biochemical network reactions in a case of biotechnological production tasks is a separate class of tasks of nonlinear differential equation

systems that can be described by a limited number of equation types and solved with the help of a precondition of the steady state existence.

Specialized software products with user interface suitable for solving biochemical tasks allow for more effective collaboration among modellers, biologists, and industrialists, compared with the universal modelling tools.

The theory of metabolic control analysis is widely used in searching the most appropriate combinations of the best variable parameters in bioprocess design tasks.

Among the existing procedures or methods there is still no procedure, which covers the full cycle of biochemical network design from the choice of criterion through to the readiness for industrial tests, by analyzing all possible combinations of optimization parameters to find and evaluate the most appropriate ones for the process.

### **3. OPTIMIZATION PROCEDURE ACTIONS AND THEIR SEQUENCE**

Following the developed steady state optimization procedure of biochemical network computer models it is possible to make effective use of resources during optimization, since it allows avoiding of unnecessary optimizations or terminating the ongoing optimization process if it is not useful. The basic principle of procedure is as follows: with every next step of the procedure to obtain the maximum possible amount of information with the least possible number of actions. In this case the information means information on either prospective or unpromising solutions of the optimization task.

The sequence of effective optimization actions in various biochemical processes can be described with steady state optimization procedure of biochemical network computer models (Mozga, Stalidzans, 2011a). Consequently, it is considered to be universal in relation to the research organism and its model, since in this respect the procedure has no limitations. The procedure in general is seen in Figure 1, and schematic representations of each step and general explanations on the steps are given in the subsequent chapter. The detailed description of the procedure lists not only the roles of various expert groups in implementing every step to avoid duplication, but also includes analysis of inappropriate or impossible variants for biology, production, or economic circumstances.

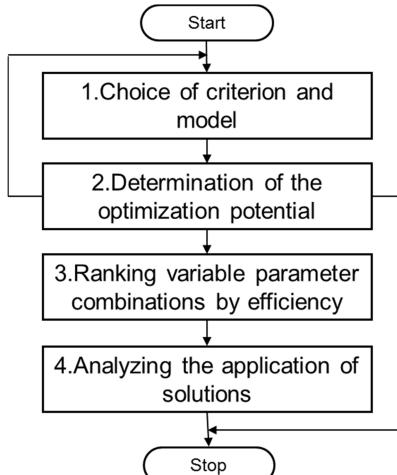


Figure 1. Sequence of optimization procedure actions

In the result of the author's research, the involved experts are divided into three groups – biologists, modellers, and industrialists. The involved expert groups relatively are attributed to different fields of expertise.

**Biologists** are familiar with mechanisms of biochemical processes at the molecular level and are able to assess the environmental parameters according to which organisms providing biotechnological process are capable of operating and what kind of changes in biochemical processes of organisms are feasible with technologies available today. Biologists must form concepts about the approximate cost of technologies.

**Modellers** have basic understanding of operating the biochemical process models, systems theory, automatic control theory, and good knowledge of optimization issues. Modellers must be competent in the available software product range to choose the most appropriate solution path for each task.

The term **industrialist** in this research refers to experts of biotechnology-oriented companies: 1) company managers and economists who are able to foresee prices of substrates and the product, and market conditions to determine economic indicators of the biotechnological process optimization, and 2) biotechnological equipment engineers who are able to assess whether the optimization process is feasible on the available facilities, or the cost of the necessary modifications.

Optimization procedure developed by the author describes actions that initially include choice of criterion and model. Namely, choosing the reactions and conditions under which to complete optimization. It results in modified numerical

values of the optimization parameters, which, through analyzing the application of solutions, may be given to industrialists for industrial testing. In the case when prolongation of optimization actions could bear the loss, the ongoing optimization is suspended and a different biochemical process is found.

### **Choice of criterion and model**

The first stage of the procedure intended for gaining overall knowledge of the process begins with choice of criterion and model – what is the business task, what kind of biochemical process is chosen, which criteria have been set. Such information is usually provided by industrialists. Then models are viewed for possible implementation with the set criteria, which in turn is done by a modeller. This stage is based on the criterion set by industrialist according to the product they want to produce, and by making economic calculations establish what the minimum criterion increase is to achieve the desired business results (see Figure 2).

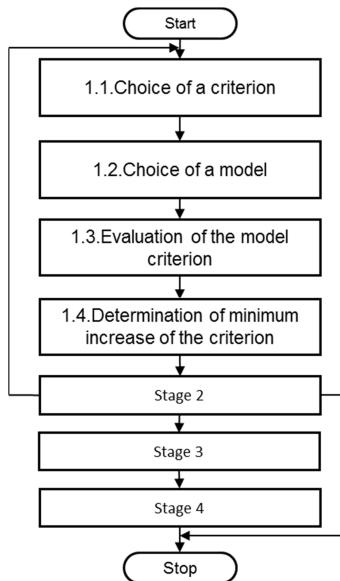
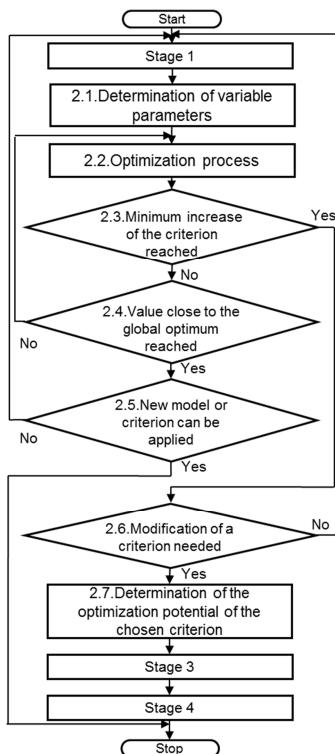


Figure 2. **Choice of criterion and model**

### **Determination of the optimization potential**

The second stage of the procedure provides for the detailed knowledge of the selected model behaviour (see Figure 3). At the beginning of the stage in collaboration with biologists, all model parameters that can be modified at the current microbiological technology development stage are selected. Further

optimizations will be carried out with the respective parameters. Then the optimization process follows, during which a number of tests are conducted to see if the chosen model with the specific modified reactions can achieve the minimum criterion increase established at the previous stage. In the case of positive response optimization is continued, otherwise the model or criterion must be changed.

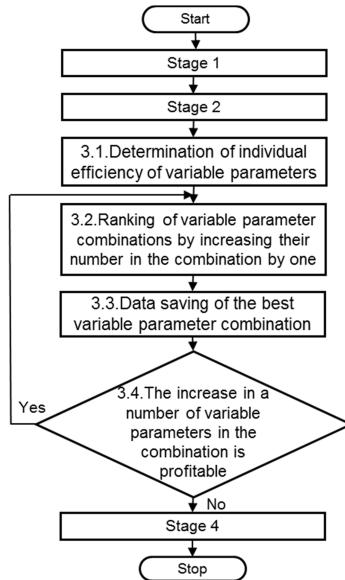


**Figure 3. Determination of the optimization potential**

#### **Ranking variable parameter combinations by efficiency**

The task of the third stage of the procedure is to find the most effective variable parameter combinations in a case of specific number of variable parameters. The efficiency is understood as the ability of variable parameters or their combinations to improve the optimization criterion value. Greater increase in the optimization criterion corresponds to a higher efficiency. It is assumed that implementation of variable parameter concentration changes in every reaction at genome level costs equally (in the procedure more precise sums can also be taken into account). At the same time it is clear that the effect of various reactions and

their combinations on the optimization criterion is very different. Therefore, the combination of variable parameters with the minimum possible number of variable parameters must be found, which could use most of the total potential of the researched process already set in step 2.7. “Determination of the optimization potential of the chosen criterion” (see Figure 4).



**Figure 4. Ranking variable parameter combinations by efficiency**

The result of applying method for evaluating the remaining range of the optimization potential can be assessed also in Figure 5: optimization was stopped at seven variable enzyme concentration coefficients (VECC) in a combination as the remaining range of optimization potential was not sufficient for gaining economic benefit in biotechnological production process. As a result, it was found that it is not worth analyzing 16,384 combinations (a case of searching all combinations), thus considerably reducing the amount of feasible work and time consumed.

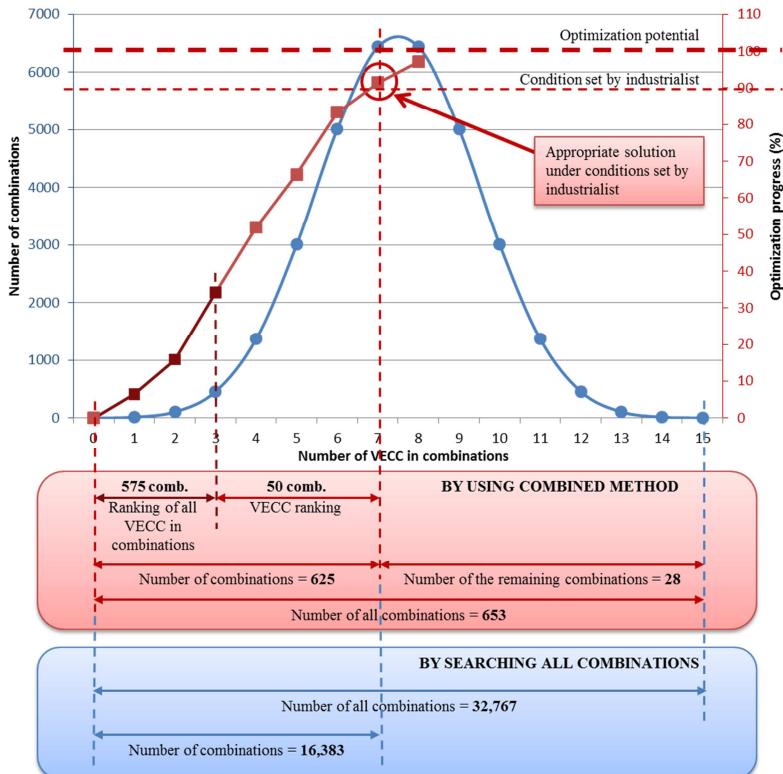


Figure 5. Application of method for evaluating the remaining range of the optimization potential for reducing the number of combinations to be analyzed

### Analyzing the application of solutions

Analyzing the application of the found solutions and experimental testing of the final obtained parameter results of the optimization process into production is carried out in Stage 4 of the procedure (see Figure 6). During Stage 4 it is verified, if the found solution not only shows the existence of a steady state, but also is sufficiently stable for implementation with the production facilities at a reasonable cost. In addition to this, it must be verified if the steady state parameters (concentration of substances, temperature, *pH* values, etc.) are feasible.

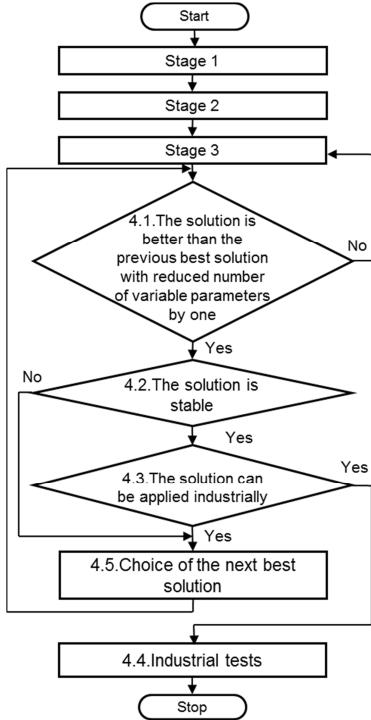


Figure 6. Analyzing the application of the found solutions

#### 4. ANALYSIS OF THE CRITERION VALUE CONVERGENCE DURING STEADY STATE OPTIMIZATION EXPERIMENT

Numerical optimization during solving of the application example appears as the most time-consuming process with ambiguous results due to the fact that the numerical methods only converge to the global optimum, the value of which is unknown. Duration of the optimization experiment can range from a few minutes up to 60-90 hours for a model with a few tens of reactions (Nikolaev, 2010). In practice it means that during optimization it is not clear at what point to stop the optimization process considering that the substantial improvement of the criterion value is not expected. Since there are many optimization methods available, and each of them tend to have parameters which, if changed, affect their efficiency, the application of methods, the rate and stability of their convergence are very important factors influencing efficiency of the optimization procedure.

## Convergence analysis tool ConvAn

It is concluded based on the experimental data described in the chapters of PhD thesis that the convergence of the optimization criterion is difficult to predict and the tools for its statistical analysis are required. Within the framework of the PhD thesis the concept of a tool was developed and functional requirements were put forward. In collaboration with Andrejs Kostromins a software tool ConvAn (Kostromins et al., 2012) with the following functionality was developed:

- optimization data storage, editing, deleting;
- registration of optimization methods;
- profiling based on the optimization conditions;
- comparison of optimization methods on the same conditions;
- analysis of a specific method under various conditions.

The name of convergence analysis tool ConvAn is made of English words *Convergence analysis*. ConvAn tool is an open-access software (<http://www.biosystems.lv/convan>) that runs on Windows and Linux operating systems (see Figure 7).

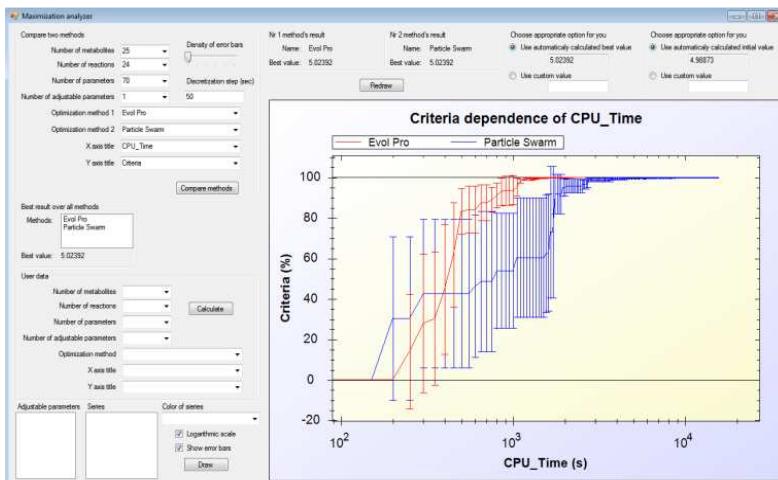


Figure 7. Comparison of ConvAn tool methods in maximization task

Input parameters of convergence analysis tool (ConvAn) are a two-vector matrix where one vector is the optimization time and the second vector is the best criterion value of the solution at that moment. The data on the model (number of species, number of reactions, number of compartments), and the parameters of the optimization task (number of variable parameters, variable parameter range / number of possible values) are entered as additional data.

ConvAn output parameters are data on the statistical evaluation of the convergence rate in the specific model for the applied optimization method. These data can be used for analyzing convergence dynamics of other optimization experiments. Also the suitability of the optimization method and its parameters can be evaluated with regard to solving the particular task.

## CONCLUSIONS

### The main PhD thesis results

A steady state optimization procedure of biochemical network computer models is developed for improvement of biotechnological production starting from the setting of optimization criteria up to justified recommendations for production experiments.

1. Optimization technology and procedure of biochemical network computer models are analyzed.

Each optimization procedure of biochemical network computer models proposed by other authors contains only part of the actions required to reach the recommendation for the best solution from the approach to the task at biotechnological production improvement level.

There is no optimization procedure of biochemical network computer models for biotechnological production needs, which takes into account the integrated interdisciplinary collaboration that would allow us to consider the timely suggestions or limitations provided by various expert groups, and to discard solutions unacceptable by one of the expert groups or the non-perspective solutions.

2. Steady state optimization procedure of biochemical network computer models is developed.

The developed steady state optimization procedure of biochemical network computer models consists of four successive stages: choice of criterion and model, determination of the optimization potential, ranking variable parameter combinations by efficiency and analyzing the application of solutions. The sequence of actions is chosen by the principle: each subsequent step in the procedure is the best by criterion a maximum amount of the new information by executing a minimum of actions.

3. Groups of experts involved in the optimization procedure interdisciplinary collaboration are determined and integrated collaboration references are developed.

Three groups of experts are involved in the optimization procedure: biologists, modellers, and industrialists (various representatives of a production

company). The developed biochemical network optimization procedure incorporates references both on groups of experts involved in making various necessary decisions during execution of the procedure and their tasks.

4. The possibilities are determined for eliminating non-perspective optimization parameter combinations in order to accelerate analysis of admissible solutions.

Method for evaluating the remaining range of the optimization potential is developed for discarding non-perspective combinations of variable parameters to be analyzed in cases when there is no reason for analyzing the remaining potential improvement due to economic reasons.

5. Factors influencing dynamics of the steady state optimization criterion of the biochemical network models are determined.

In solving steady state optimization tasks of biochemical network models the most of the popularity is gained by global stochastic numerical optimization algorithms that are universally applicable regardless of the model characteristics and the type of nonlinearity.

The convergence rate of stochastic global optimization methods and the reaching of global optimum depend both on the number of variable parameters and the chosen variable parameters in the case of a constant number of variable parameters, and on the chosen optimization methods.

6. A tool for dynamics data processing and analysis of the best steady state optimization criterion value of the biochemical network models is developed.

A tool ConvAn is developed for evaluating and forecasting the optimization convergence rate. Multiple optimizations with identical initial conditions are evaluated by statistical methods. The convergence rate based on the analysis is predictable for an equivalent model using the same method. ConvAn can be used in evaluating and forecasting the convergence rate of the optimization criterion both in the cases of parameter evaluation and design tasks.

7. Operation of procedure is demonstrated by justifying recommendations for industrial production of ethanol from glucose using the baker's yeast (*Saccharomyces cerevisiae*) glycolysis model.

The coefficients for changing enzyme concentrations in seven reactions allowing for changes in the range 0.01 to 11.0 are developed for the baker's yeast (*Saccharomyces cerevisiae*) glycolysis model created by Hynne after the execution of the optimization procedure (Hynne et al., 2001). Manipulable reactions and their coefficients (in parentheses) are as follows: Glucose Uptake (11.0), Hexokinase (2.73), Phosphofructokinase (2.02), Pyruvate Decarboxylase (5.93), Alcohol Dehydrogenase (11.0), Storage (0.01), and

ATP consumption (11.0). The model predicts that after introducing changes the Ethanol Flow increases from 0.80 to 2.05 mmol/min, but Glucose Uptake increases from 0.83 to 1.12 mmol/min, and the following side products are reduced: glycerol from 0.085 mmol/min to 0.028 mmol/min, acetaldehyde from 0.065 mmol/min to 0.021 mmol/min, and cyanide from 0.020 mmol/min to 0.007 mmol/min. In the result, both the product yield of substrate unit is improved from 0.97 to 1.83, and ethanol yield is increased 2.6 times.

### **Conclusions and development prospects**

While the number of biochemical models and their size increases, the model optimization procedures continue to develop, where mostly the stochastic global optimization methods are applied due to their relatively quick and safe convergence properties. These methods may be combined also in hybrid methods, where there is a shift from stochastic global methods to local methods and vice versa.

The dynamics of biochemical network reactions in a case of biotechnological production tasks is a separate class of tasks of nonlinear differential equation systems that can be described by a limited number of equation types and solved with the help of a precondition of the steady state existence.

The choice of a suitable optimization criterion in a case of the design task is made iteratively in collaboration with biologists, designers, and industrialists, and depends not only on economic and technological factors but also on the model size and its characteristics.

The ranking of variable model parameters in a case of the steady state optimization of biochemical network models is lagging behind by efficiency from the ranking of combinations of variable model parameters.

The developed combined method for ranking a large number of variable parameters offers both to use analysis of all combinations, if the number of parameters in combination is small, and to move to the ranking of parameters by searching the best parameter to be involved in the next combination from the remaining parameters. Such approach can not guarantee finding of the best combinations for a larger number of parameters, but the result can be achieved within a reasonable time. In the case when a number of all combinations is small or there are enough resources for testing all possible combinations, all combinations must be tested.

The convergence rate of stochastic global optimization methods and reaching the value close to the global optimum under conditions of a constant model, initial

conditions and optimization method settings may depend on the number of variable parameters, their properties, and the chosen optimization method.

In a case of the total optimization potential, changing 15 enzyme concentrations by using the optimization procedure, there is a successful increase in ethanol productivity from 0.80 mmol/min to 2.27 mmol/min and an improvement of product yield of substrate unit from 0.97 to 1.97, and a reduction of the following side products: glycerol from 0.085 mmol/min to 0.007 mmol/min, acetaldehyde from 0.065 mmol/min to 0.005 mmol/min, and cyanide from 0.020 mmol/min to 0.002 mmol/min.

Several directions can be highlighted as **future development prospects**:

1. Develop criteria for terminating the optimization experiment using global stochastic methods based on the statistical analysis of convergence dynamics of several parallel optimizations.
2. To analyze variable parameter combinations a software product must be developed that would automatically perform testing of all possible combinations and their ranking by efficiency for various number of variable parameters using optimization criteria for terminating the experiment and discarding non-perspective solutions based on the curves of average convergence dynamics.
3. Look for the opportunities to test the developed recommendations for changing enzyme concentrations in biological experiments.
4. Compare the optimization methods available in open-access software products by the best solution of convergence dynamics using ConvAn tool.

## LITERATŪRAS SARAKSTS BIBLIOGRAPHY

1. Acerenza L. (1993) Metabolic control design. *Journal of theoretical biology*, 165(1), pp. 63-85.
2. Acerenza L. (2000) On the Universality of The Universal Method. In a. Cornish-Bowden & C. M. L (Eds.), *Technological and medical implications of metabolic control analysis*, pp. 33-37.
3. Acerenza L., Ortega F. (2007) Modular metabolic control analysis of large responses. *The FEBS journal*, 274(1), pp. 188-201.
4. Achard F., Vaysseix G., Barillot E. (2001) XML, bioinformatics and data integration. *Bioinformatics* 17:115-125.
5. Alberghia L., Hohmann S., Westerhoff H.V. (2005) Systems Biology: necessary developments and trends. In book: Alberghina L., Westerhoff H.V. (eds.) *Systems Biology, Definitions and Perspectives*. Springer, pp. 390-402.
6. Ali, M. M., Storey, C., & Torn, A. (1997) Application of Stochastic Global Optimization Algorithms to Practical Problems. *Journal of Optimization Theory and Applications*, 95(3), pp. 545-563.
7. Alm E., Arkin A.P. (2003) Biological networks. *Curr Opin Struct Biol* 13:193-202.
8. Assmus H.E., Cho K.H., Herwig R., Wolkenhauer O. (2006) Dynamics of biological systems: role of systems biology in medical research. *Expert Rev. Mol. Diagn.* 6(6), pp. 891-902.
9. Arkin A.P. (2001) Simulac and deduce. [tiešsaiste] [skatīts 24.10.2008.] Pieejams: <http://gobi.lbl.gov/~aparkin/Stuff/Software.html>.
10. Augen J. (2001) Information technology to the rescue! *Nat Biotechnol* 19:BE39-BE40
11. Bäck T., Schwefel H.-P. (1997) An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*. 1, pp. 1-23.
12. Bäck T., Fogel D.B., Michalewicz Z. (1997) *Handbook of evolutionary computation*. Oxford: IOP Publishing/Oxford University Press.
13. Baker, S. M., Schallau, K., & Junker, B. H. (2010) Comparison of different algorithms for simultaneous estimation of multiple parameters in kinetic metabolic models. *Journal of integrative bioinformatics*, 7(3), pp. 1-9. doi: 10.2390/biecoll-jib-2010-133.
14. Bakker B., Helfert S., Estévez A.M., Michels P., Clayton C. (2001) Roles of triosephosphate isomerase and aerobic metabolism in *Trypanosoma brucei*. *Biochem J*, 357(Pt 1):1, pp. 17-25.
15. Balsa-Canto, E., Alonso, A. A., & Banga, J. R. (1998) Dynamic optimization of bioprocesses : deterministic and stochastic strategies. In C. Skjoldebremd (Eds.), *ACoFop IV (Automatic Control of Food and Biological Processes)* (Vol. 2, pp. 1-6). Goteborg, Sweden.

16. Balsa-Canto, E., Peifer, M., Banga, J. R., Timmer, J., & Fleck, C. (2008) Hybrid optimization method with general switching strategy for parameter estimation. *BMC systems biology*, 2, 26. doi: 10.1186/1752-0509-2-26.
17. Balsa-Canto, E., Alonso, A. a, Banga, J. R. (2010) An iterative identification procedure for dynamic modeling of biochemical networks. *BMC systems biology*, 4, 11. doi: 10.1186/1752-0509-4-11.
18. Banga, J. R. (2008) Optimization in computational systems biology. *BMC systems biology*, 2, 47. doi: 10.1186/1752-0509-2-47.
19. Bell M., Pike M.C. (1966) Remark on algorithm 178 direct search. *Communications of the Association for Computing Machinery*. 9. pp. 684 - 685.
20. Bhattacharyya B.C., Banerjee R. (2007) Environmental Biotechnology, Oxford University Press, New Delhi, India, 340 p.
21. Biotechnical Center (BTC). Biopopulārā tehnoloģija. [tiešsaiste] [skatīts 25.09.2008.] Pieejams: <http://www.bioreactors.net/lat/biotechnology.html>
22. Bonnans J.F., Gilbert J.C., Lemarechal C., Sagastizabal C. (2006) Numerical optimization: theoretical and practical aspects, Universitext Series, Springer-Verlag Berlin Heidelberg. 490 p.
23. Bosak, J., Bray, T. (1999) XML and the second-generation Web. *Scientific American*, 280 (5), pp. 79-83.
24. Bray T., Paoli J., Sperberg-McQueen C.M., Maler E. (2000) Extensible markup language (XML) 1.0 Second Edition:<http://www.w3.org/TR/1998/REC-xml-19980210> [tiešsaiste] [skatīts 24.10.2008.]
25. Brent P.R. (1973) A new algorithm for minimizing a function of several variables without calculating derivatives. In Algorithms for minimization without derivatives, (Englewood Cliffs, NJ: Prentice-Hall, Inc.). pp. 117 - 167.
26. Bro, C., Regenberg, B., Foerster, J., & Nielsen, J. (2006) In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metabolic Engineering*, 8, pp. 102-111.
27. Brugård J. (2008) Integrative Genomics and Chronic Disease Phenotypes: modelling and simulation tools for clinicians. BioBridge project document. pp. 17.
28. Bunge M.A. (1967) Scientific Research. Springer, Berlin, Germany.
29. Burns J.A. (1971) Studies on complex enzyme systems. University of Edinburgh.
30. Butler D. (1999) Computing 2010: from black holes to biology. *Nature* 402:C67-C70.
31. Capra F. (1996) The Web of Life: A new scientific understanding of living systems. Anchor Books, New York.
32. Castle Rock Renewable fuels. Ethanol Production Process. [tiešsaiste] [skatīts 07.12.2011.] Pieejams: <http://www.castlerockethanol.com/contentDetail.asp?id=8754>

33. Chang Y., Sahinidis N. (2005) Optimization of metabolic pathways under stability considerations. *Computers & Chemical Engineering*, 29(3), pp. 467-479. doi: 10.1016/j.compchemeng.2004.08.013.
34. Corana A., Marchesi M., Martini C., and Ridella S. (1987) Minimizing multimodal functions of continuous variables with the "simulated annealing" algorithm. *ACM Trans. Math. Softw.* 13. pp. 262 - 280.
35. Cornish-Bowden A. (2004) Fundamentals of Enzyme Kinetics, 3rd edn., 344 p. Portland Press, London.
36. Crabtree B., Newsholme E.A. (1985) A quantitative approach to metabolic control. *Curr Top Cell Regul* 25:21-76.
37. Dorf, R., & Bishop, R. (2005) Modern Control Systems (10th ed., 634 p.). Pearson Education.
38. EC (2003) Directive 2003/30/EC of the European Parliament and of the Council of 8 May 2003 on the promotion of the use of biofuels or other renewable fuels for transport. *Official J Eur Union L* 123:42–46.
39. Ekmanis J. (2008) Latvijas ZA - Par LZA.[tiešsaiste] [skatīts 25.09.2008.] Pieejams: [http://www.lza.lv/index.php?option=com\\_content&task=view&id=47&Itemid=56](http://www.lza.lv/index.php?option=com_content&task=view&id=47&Itemid=56)
40. Enari T-M. (1999) From beer to molecular biology: the evolution of industrial biotechnology. Nurnberg: Fachverlag Hans Carl, 120 p.
41. Ennos R. (2000) Statistical and data handling skills in biology. Pearson Education, 132.
42. Fell D.A. (2005) Metabolic Control Analysis. In: Alberghina L., Westerhoff H.V. (eds.) *Systems Biology, Definitions and Perspectives*, Springer-Verlag Berlin Heidelberg, Germany, pp. 69-80.
43. Fell D.A., Sauro H.M. (1985) Metabolic Control Analysis: Additional relationships between elasticities and control coefficients. *Eur J Biochem* 148:555-561.
44. Fink C.C., Slepchenko B., Moraru I.I., Watras J., Schaff J.C. Loew L.M. (2000) An image-based model of calcium waves in differentiated neuroblastoma cells. *Biophys J* 79:163-83.
45. Finney A., Hucka M. (2003) Systems biology markup language: level 2 and beyond. *Biochem Soc Trans.*, 31:1472–1473.
46. Floudas, C. A. (2000) Deterministic global optimization: Theory, methods and applications. *Computers & Mathematics with Applications* (Vol. 40, 415 p.). Dordrecht/Boston/London: Kluwer Academic Publishers. doi: 10.1016/S0898-1221(00)90165-2.
47. Fogel D.B., Fogel L.J., Atmar J.W. (1992) Meta-evolutionary programming. 25th Asiloma Conference on Signals, Systems and Computers. IEEE Computer Society, Asilomar. pp. 540-545.
48. Fraser S.E., Harland R.M. (2000) The molecular metamorphosis of experimental embryology. *Cell* 100:41-55.

49. Funahashi A., Tanimura N., Morohashi M., Kitano H. (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. BioSilico 1:159-162.
50. Funahashi A., Tanimura N., Morohashi M., Kitano H. (2004) CellDesigner; <http://www.systems-biology.org/002/> [tiešsaiste] [skatīts 24.10.2008.]
51. Galazzo J. L., Bailey J. E. (1990) Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae*. Enzyme and Microbial Technology, 12(3), pp. 162-172.
52. Gill P.E., Murray W., and Wright M.H. (1981) Practical Optimization. London, Academic Press.
53. Gillespie D.T. (1977) Exact stochastic simulation of coupled chemical-reactions. *J Phys Chem* 81:2340-2361.
54. Gillespie D.T., Petzold L.R. (2003) Improved leap-size selection for accelerated stochastic simulation. *J Chem Phys* 119:8229-8234.
55. Goldfeld S. M., Quant R. E., and Trotter H. F. (1966) Maximisation by quadratic hill-climbing. *Econometrica* . 34. pp. 541-555.
56. Goldstein B., Faeber J.R., Hlavacek W.S., Blinov M.L., Redondo A., Wolfs C (2002) Modeling the early signaling events mediated by FceRI. *Mol Immunol* 137:1-7.
57. Grunde-Zeiferts, U., Mozga, I., Žukova, T., & Stalidzans, E. (2006) Terapijas modelēšana apvienojot sistēmbioloģijas un automātiskās vadības teorijas metodes. *Animals. Health. Food hygiene.* pp. 70-74. Jelgava.
58. Hanisch D., Zimmer R. and Lengauer T. (2002) ProML-the protein markup language for specification of protein sequences, structures and families. In *Silico Biol.* 2, pp. 313-24.
59. Hartwell L.H., Hopfield J.J., Leibler S., Murray A.W. (1999) From molecular to modular cell biology. *Nature* 402:C47-C52.
60. Hatzimanikatis V. (1999) Nonlinear metabolic control analysis. *Metabolic engineering*, 1(1), pp. 75-87.
61. Heinrich R., Rapoport T.A. (1974) A linear steady-state treatment of enzymatic chains; general properties, control and effector strength. *Eur J Biochem* 42:89-95.
62. Hirmajer, T., Balsa-Canto, E., & Banga, J. R. (2009) DOTcpSB, a software toolbox for dynamic optimization in systems biology. *BMC bioinformatics*, 10, 199. doi: 10.1186/1471-2105-10-199.
63. Hisiger, S., Jolicoeur, M., Herrera, F., Deraz, S. F., Karlsson, E. N., M, M., et al. (2005) Dynamic optimization of bioprocesses: Efficient and robust numerical strategies. *Journal of Biotechnology*, 117(4), pp. 407-419.
64. Hoefnagel M.H., Starrenburg M.J., Martens D.E., Hugenholtz J., Kleerebezem M., Van Swam I.I., Bongers R., Westerhoff H.V., Snoep J.L. (2002) Time dependent responses of glycolytic intermediates in a detailed glycolytic model of *Lactococcus lactis* during glucose run-out experiments. *Mol Biol Rep*, 29(1-2):1, pp. 57-61.

65. Hood L. (1998) Systems biology: New opportunities arising from genomics, proteomics, and beyond. *Exp Hematol* 26:681.
66. Hooke R., Jeeves T. A. (1961) "Direct search" solution of numerical and statistical problems. *Journal of the Association for Computing Machinery*. 8. pp. 212-229.
67. Hoops S., Sahle S., Gauges R., Lee C., Pahle J., Simus N., Singhal M., Xu L., Mendes P. and Kummer U. (2006) COPASI — a COmplex PAthway SImulator. *Bioinformatics*, 22, pp. 3067-74.
68. Hucka M., Finney A., Sauro H.M., Bolouri H. (2001) Systems Biology Markup Language (SBML) Level 1: Structures and facilities for basic model definitions; <http://www.sbml.org/> [tiešsaiste] [skatīts 24.10.2008.]
69. Hucka M., Finney A., Sauro H.M., Bolouri H., Doyle J.C., Kitano H., Arkin A.P., Bornstein B.J., Bray D., Cornish-Bowden A., Cuellar A.A., Dronov S., Gilles E.D., Ginkel M., Gor V., Gryanin I.I., Hedley W.J., Hodgman T.C., Hofmeyr J.H., Huter P.J., Juty N.S., Kasberger J.L., Kremling A., Kummer U., Le Novere N., Loew L.M., Lucio d., Mendes P., Minch E., Mjosness E.D., Nakayama Y., Nelson M.R., Nielsen P.F., Sakurada T., Schaff J.C., Shapiro B.E., Shimizu T.S., Spence H.D., Stelling J., Takahashi K., Tomita M., Wagner J., Wang J. (2003a) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. 1. *Bioinformatics*. 19:513–523.
70. Hucka M., Finney A., Sauro H.M., Bolouri H., Doyle J.C., Kitano H., Arkin A.P., Bornstein B.J., Bray D., Cornish-Bowden A., Cuellar A.A., Dronov S., Gilles E.D., Ginkel M., Gor V., Gryanin I.I., Hedley W.J., Hodgman T.C., Hofmeyr J.H., Huter P.J., Juty N.S., Kasberger J.L., Kremling A., Kummer U., Le Novere N., Loew L.M., Lucio d., Mendes P., Minch E., Mjosness E.D., Nakayama Y., Nelson M.R., Nielsen P.F., Sakurada T., Schaff J.C., Shapiro B.E., Shimizu T.S., Spence H.D., Stelling J., Takahashi K., Tomita M., Wagner J., Wang J. (2003b) The Systems Biology Markup Language (SBML): a medium for representation and exchange of biochemical network models. 2. *Bioinformatics* 19:524-531.
71. Hynne F., Danø S., Sørensen P.G. (2001) Full-scale model of glycolysis in *Saccharomyces cerevisiae*. *Biophys Chem*, 94(1-2):1, pp. 21-63.
72. Ideker T., Galitski T., Hood L. (2001) A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* 2:343-372.
73. Ingalls, B. P., Iglesias, P. (2010) A Primer on Control Engineering. *Control Theory and Systems Biology*, pp. 1-28.
74. Jain E., Bairoch A., Duvaud S., Phan I., Redaschi N., Suzek B.E., Martin M.J., McGarvey P., Gasteiger E. (2009) Infrastructure for the life sciences: design and implementation of the UniProt website, *BMC Bioinformatics*, 10:136.
75. Kacser H. (1957) Appendix: Some physic-chemical aspects of biological organization. In Waddington C.H.(ed) *The strategy of the genes: A*

- discussion of some aspects of theoretical biology. George Allen and Unwin Ltd, London, pp. 191-249.
- 76. Kacser H., Acerenza L. (1993) A universal method for achieving increases in metabolite production. European journal of biochemistry / FEBS, 216(2), pp. 361-7.
  - 77. Kacser H., Burns J.A. (1973) The control of flux. Symp Soc Exp Biol 27:65-104. Reprinted in Biochem Soc Trans (1995) 23:341-366.
  - 78. Kaupe. (1963) Algorithm 178 direct search. Communications of the Association of Computing Machinery. 6. pp. 313 - 314.
  - 79. Keasling J.D. (2010) Manufacturing molecules through metabolic engineering. Science (New York, N.Y.), 330(6009), pp. 1355-8.
  - 80. Kennedy J., Eberhart R. (1995) Particle Swarm Optimization. Proceedings of the Fourth IEEE International Conference on Neural Networks, Perth, Australia. pp. 1942-1948.
  - 81. Kirkpatrick S., Gelatt J., C.D., Vecchi M. P. (1983) Optimization by simulated annealing. Science. 220. pp. 671-680.
  - 82. Kitano, H. (2000) Perspective on Systems Biology. New Generation Computing, 18, pp. 199-216.
  - 83. Kitano H. (2001) Foundations of systems biology. MIT Press, Cambridge, MA.
  - 84. Kitano H. (2002a) Computational systems biology. Nature 420:206-210.
  - 85. Kitano H. (2002b) Systems Biology Markup Language Standard. 2002. pp. 5.
  - 86. Klipp E., Herwig R., Kowald A., Wierling C. and Lehrach H. (2005) Systems Biology in Practice, Concepts, Implementation and Application, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 467 p.
  - 87. Kondili E.M., Kaldellis J.K. (2007) Biofuel implementation in East Europe: Current status and future prospects. Renewable and Sustainable Energy Reviews 11 , pp. 2137-2151.
  - 88. Kostromins A., Mozga I., Stalidzans E. (2012) ConvAn: a convergence analyzing tool for optimization of biochemical networks. Biosystems, 108(1-3), pp. 73-77. Elsevier. doi:10.1016/j.biosystems.2011.12.004.
  - 89. Lagzdina L., Mozga I., Stalidzans E. (2008) Algorithm for bioethanol yield optimization in *Saccharomyces cerevisiae* glycolysis using COPASI. Applied Information and Communication Technologies, pp. 95-101. Jelgava.
  - 90. Lee, F. C., Pandu Rangaiah, G., Lee, D.-Y. (2010) Modeling and optimization of a multi-product biosynthesis factory for multiple objectives. Metabolic engineering, 12(3), pp. 251-67. Elsevier. doi: 10.1016/j.ymben.2009.12.003.
  - 91. Levenberg K. (1944) A method for the solution of certain nonlinear problems in least squares. Quart. Appl. Math. 2. pp. 164-168.

92. Li A., Antizar-Ladislao B., Khraisheh M. (2007) Bioconversion of municipal solid waste to glucose for bio-ethanol production. *Bioprocess Biosyst Eng* 30, pp. 189-196.
93. Li C., Donizelli M., Rodriguez N., Dharuri H., Endler L., Chelliah V., Li L., He E., Henry A., Stefan M.I., Snoep J.L., Hucka M., Le Novère N. and Laibe C., (2010) BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology* 2010, 4:92
94. Liao Y.M., Ghanadan H. (2002) The chemical markup language. *Anal. Chem.* 74, 389A-390A.
95. Luisi P. L., Ferri F., Stano P. (2006) Approaches to semi-synthetic minimal cells: a review. *Die Naturwissenschaften*, 93(1), pp. 1-13.
96. Magnus J. B., Oldiges M., Takors R. (2009) The identification of enzyme targets for the optimization of a valine producing *Corynebacterium glutamicum* strain using a kinetic model. *Biotechnology progress*, 25(3), pp. 754–762. Wiley Online Library.
97. Maier W.L. (1991) A Fast Pseudo Random Number Generator. *Dr. Dobb's Journal*. May. pp. 152-157.
98. Maileret, L., Gouzé, J.-L., & Bernard, O. (2005) Nonlinear control for algae growth models in the chemostat. *Bioprocess and biosystems engineering*, 27(5), pp. 319-27.
99. Maiwald T., Timmer J. (2008) Dynamical Modeling and Multi-Experiment Fitting with PottersWheel. *Bioinformatics*, 24(18):2037-2043.
100. Marquardt D.W. (1963) An algorithm for least squares estimation of nonlinear parameters. *SIAM Journal*. 11. pp. 431-441.
101. Matsumoto M., Nishimura T. (1998) Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans. on Modeling and Computer Simulations*. 8. pp. 3-30.
102. Mauch K., Buziol S., Schmid J., Reuss M. (2001) Computer-Aided Design of Metabolic Networks. *AIChE Symposium Series*, pp. 82-91.
103. Mendes P. (1993) GEPASI: a software package for modeling the dynamics, steady states and control of biochemical and other systems. *Comput Appl Biosci* 9, pp. 563-71.
104. Mendes P. (1997) Biochemistry by numbers: Simulation of biochemical pathways with Gepasi 3. *Trends Biochem. Sci.* 22, pp. 361-363.
105. Mendes, P., Kell, D. (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics* (Oxford, England), 14(10), pp. 869-83.
106. Mesarovic M.D. (1968) System theory and biology. *Proceedings of the 3rd Systems Symposium at Case Institute of Technology*. Springer-Verlag, Berlin, New York.
107. Michalewicz Z. (1994) Genetic algorithms + data structures = evolution programs. 3rd Edition. Springer-Verlag, Berlin.

108. Mitchell M. (1995) An Introduction to Genetic Algorithms. MIT Press, Boston.
109. Moles, C. G., Mendes, P., Banga, J. R. (2003) Parameter estimation in biochemical pathways: a comparison of global optimization methods. (C. Skjoldebremd & G. Trystrom, Eds.) *Genome Research*, 13(11), pp. 2467-2474. Gotheborg, Sweden.
110. Mozga, I., Grunde-Zeiferts, U., Sudars, S., & Stalidzans, E. (2007) Life cycles and competition in modelling of artificial and biological control system. European Simulation Multiconference, pp. 533-536. Malta: EUROSIS.
111. Mozga I., Kostromins A., Stalidzans E. (2011) Forecast of numerical optimization progress of biochemical networks. 10th International Scientific Conference „Engineering for rural development”, pp. 103-108.
112. Mozga, I., Stalidzans, E. (2007) Application of dynamic models of glycolysis developing control system. *Engineering for Rural Development*, pp. 119-123. Jelgava.
113. Mozga I., Stalidzans E. (2011a) Optimization protocol of biochemical networks for effective collaboration between industry representatives, biologists and modellers. 9th International Industrial Simulation Conference (ISC'2011), pp. 91-96.
114. Mozga I., Stalidzans E. (2011b) Convergence dynamics of biochemical pathway steady state stochastic global optimization. 2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI), pp. 231-235. IEEE.
115. Mozga I., Stalidzans E. (2011c) Convergence Dynamics of Biochemical Models To The Global Optimum, 3rd IEEE International Conference on e-Health and Bioengineering, EHB 2011 – Proceedings, pp. 227-230.
116. Nash S. G. (1984) Newton-type minimization via the Lanczos method. *SIAM Journal of Numerical Analysis*. 21. pp. 770-788.
117. Nelder J. A., Mead R. (1965) A simplex method for function minimization. *Computer Journal*. 7. pp. 308-313.
118. Newsholme E.A., Start C. (1973) Regulation in Metabolism. Wiley and Sons, London.
119. Nielsen K., Sørensen P. G., Hynne F. and Busse H. -G. (1998) Sustained oscillations in glycolysis: an experimental and theoretical study of chaotic and complex periodic behavior and of quenching of simple oscillations. *Biophysical Chemistry*, Volume 72, Issues 1-2, pp. 49-62.
120. Nielsen, J. (2007) Principles of optimal metabolic network operation. *Molecular systems biology*, 3(126), 126.
121. Nikolaev E.V. (2010) The elucidation of metabolic pathways and their improvements using stable optimization of large-scale kinetic models of cellular systems. *Metabolic engineering*, 12(1), pp. 26-38. Elsevier.

122. Noble D. (2002) The rise of computational biology. *Nat Rev Mol Cell Biol* 3:460-463.
123. Olivier B.G., Snoep J.L. (2004), Web-based kinetic modelling using JWS Online, *Bioinformatics* 20, pp. 2143-2144.
124. Ortega F., Acerenza L. (1998) Optimal metabolic control design. *Journal of theoretical biology*, 191(4), pp. 439-49.
125. Palsson, B. (2000) The challenges of in silico biology Moving from a reductionist paradigm to one that views cells as systems will necessitate. *Nature Biotechnology*, 18(November), pp. 1147-1150
126. Powell M.J.D. (1964) An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* 7. pp. 155-162.
127. Pritchard L., Kell D.B. (2002) Schemes of flux control in a model of *Saccharomyces cerevisiae* glycolysis. *Eur J Biochem*, 269(16):3, pp. 894-904.
128. Reder C. (1988) Metabolic control theory: a structural approach. *J Theor Biol* 135:175-201.
129. Reiss, T. (2002) Systeme des Lebens Systembiologie (p. 34). Heidelberg.
130. Reiss, T. (2005) Workpackage 1: International benchmarking and foresight of systems biology (EUSYSBIO), 44 p. Karlsruhe.
131. Roberts, R. J. (2001) PubMed Central: The GenBank of the published literature. *Proceedings of the National Academy of Sciences* 98 (2): 381-382.
132. Rodriguez-Acosta F., Regalado C.M., Torres N.V. (1999) Non-linear optimization of biotechnological processes by stochastic algorithms: application to the maximization of the production rate of ethanol, glycerol and carbohydrates by *Saccharomyces cerevisiae*. *Journal of biotechnology*, 68(1), pp. 15-28.
133. Rodriguez-Fernandez, M., Mendes, P., Banga, J. R. (2006) A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Bio Systems*, 83(2-3), pp. 248-65.
134. Rodriguez-Prados J.-C., de Atauri P., Maury J., Ortega F., Portais J.-C., Chassagnole C., Acerenza L., et.al. (2009) In silico strategy to rationally engineer metabolite production: A case study for threonine in *Escherichia coli*. *Biotechnology and bioengineering*, 103(3), pp. 609-20.
135. Runarsson T., Yao X.. (2000) Stochastic ranking for constrained evolutionary optimization. *IEEE Transactions on Evolutionary Computation*. 4. pp. 284-294.
136. Sauro H.M. (2001) JDesigner: A simple biochemical network designer; <http://members.tripod.co.uk/sauro/biotech.htm> [tiesīsaiste] [skatīts 24.10.2008.]
137. Sauro H.M. (2003) WinScamp; [tiesīsaiste] [skatīts 24.10.2008.] Pieejams: <http://www.cds.caltech.edu/~hsauro/Scamp/scamp.htm>

138. Savageau M.A. (1969) Biochemical systems analysis.1. Some mathematical properties of rate law for component enzymatic reactions. *J Theor Biol* 25:365-366.
139. Savageau M.A. (1970) Biochemical systems analysis.3. Dynamic solutions using a power-law approximation. *J Theor Biol* 26:215.
140. Savageau M.A. (1976) Biochemical Systems Analysis: a Study of Function and Design in Molecular Biology. Addison-Wesley, Reading, Mass. 379 p.
141. Schmidt H., Jirstrand M. (2006) Systems Biology Toolbox for MATLAB: a computational platform for research in Systems Biology. *Bioinformatics*, 22(4), pp. 514-515.
142. Schuetz, R., Kuepfer, L., Sauer, U. (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Molecular systems biology*, 3(119), 119.
143. Sendín J.O.H., Exler O., Banga, J.R. (2010) Multi-objective mixed integer strategy for the optimisation of biological networks. *IET systems biology*, 4(3), pp. 236-48.
144. Small J. R., Kacser H. (1993a) Responses of metabolic systems to large changes in enzyme activities and effectors. 2. The linear treatment of branched pathways and metabolite concentrations. Assessment of the general non-linear case. *European Journal of Biochemistry*, 213(1), pp. 625-640. Wiley Online Library.
145. Small J. R., Kacser H. (1993b) Responses of metabolic systems to large changes in enzyme activities and effectors. 1. The linear treatment of unbranched chains. *European Journal of Biochemistry*, 213(1), pp. 613-624.
146. Spellman P.T., Miller M., Stewart J., Troup C., Sarkans U., Chervitz S., Bernhart D., Sherlock G., Ball C., Lepage M., Swiatek M., Marks W.L., Goncalves J., Markel S., Iordan D., Shojatalab M., Pizarro A., White J., Hubley R., Deutsch E., Senger M., Aronow B.J., Robinson A., Bassett D., Stoeckert C.J., Jr. and Brazma A. (2002) Design and implementation of microarray gene expression markup language (MAGE-ML). *Genome Biol*. 3, RESEARCH0046.
147. Stalidzans, E., Markovics, Z. (2000) Expert based model building using incidence matrix and topological model. 12th European Simulation Symposium, pp. 328-332. Hamburg.
148. Stalidzans, E. (2005). Datorvadības algoritmi daudzobjektu bioloģiskām sistēmām. Rīgas Tehniskā Universitāte.
149. Stalidzans, E., Markovics, Z. (2005) Methodology of control system development for biological systems under information insufficiency. European Modelling Simulation Symposium, pp. 169-175. Marseille.
150. Stalidzans, E., Gailums, A., Mozga, I. (2006) Life cycle of artificial and biological control system. Information Technologies for Rural Development, pp. 164-172. Jelgava.

151. Stalidzans E., Mozga I. (2006) Acquisition of dynamic patrameters of biological process from experimental data. Information Technologies for Rural Development, pp. 131-140. Jelgava.
152. Stalidzans E., Mozga I. (2006) Control oriented dynamic modelling of biological object. Biometrics and information technologies in agriculture: research and development, pp. 11-14. Kaunas.
153. Stelling J., Kremling A., Ginkel M., Bettendorf K., Gilles E. (2001) Towards a virtual biological laboratory. In: Kitano H. (ed) Foundations of systems biology. MIT Press, Cambridge, MA, pp. 189-212.
154. Stephanopoulos G., Simpson T. W. (1997) Flux amplification in complex metabolic networks. Chemical Engineering Science, 52(15), pp. 2607-2627.
155. Swann W.H. (1972) Direct search methods. Numerical methods for unconstrained optimization., W. Murray, ed. (London & New York: Academic Press). pp.13 - 28. Systems Engineering Fundamentals. Defense Acquisition University Press, 2001.
156. Teusink B, Passarge J, Reijenga CA, Esgalhado E, van der Weijden CC, Schepper M, Walsh MC, Bakker BM, van Dam K, Westerhoff HV, Snoep JL. (2000) Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. Eur J Biochem, 267(17):53, pp. 13-29.
157. Tomita M., Hashimoto K., Takahashi K. Shimizu T.S., Matsuzaki Y., Miyoshi F., Saito K., Takanida S., Yugi K., Venter J.C., Hutchison C.A., III (1999) E-CELL: software environment for whole-cell simulation. Bioinformatics 15:72-84.
158. Tomita M., Nakayama Y., Naito Y., Shimizu T., Hashimoto K., Takahashi K., Matsuzaki Y., Yugi K., Miyoshi F., Saito Y., Kuroki A., Ishida T., Iwata T., Yoneda M., Kita M., Yamada Y., Wang E., Seno S., Okayama M., Kinoshita A., Fujita Y., Matsuo R., Yanagihara T., Watari d., Ishinabe S., Miyamoto S. (2001) E-CELL; <http://www.e-cell.org/> [tiešsaiste] [skatīts 24.10.2008.]
159. Tyson J.J., Chen K., Novak B. (2001) Network dynamics and cell physiology. Nat Rev Mol Cell Biol 2:908-916.
160. Viesturs U., Simeonov I., Pencheva T., Vanags J., Petrov M., Pavlov Y., Roeva O., Ilkova T., Vishkins M., Hristozov I. (2010) Contemporary Approaches to Modelling, Optimization and Control of Biotechnological Processes. (Ed.Stoyan Tzonkov). Sofia: Prof.Marin Drinov Academic Publishing House, 240 p.
161. Vīgants A. (2008) Cilvēka bioķīmija un molekulārā bioloģija. Rīga: LU Akadēmiskais apgāds. 122. lpp.
162. Vital-Lopez, F. G., Armaou, A., Nikolaev, E. V., Maranas, C. D. (2006) A computational procedure for optimal engineering interventions using kinetic models of metabolism. Biotechnology progress, 22(6), pp. 1507-17.

163. Wanner B.L., Finney A., Hucka M., Modeling the E. coli cell: The need for computing, cooperation, and consortia. In book: Alberghina L., Westerhoff H.V. (eds.) Systems Biology, Definitions and Perspectives. Springer, 2005., pp.163-189.
164. Westerhoff H.V., Kolodkin A., Conradie R., Wilkinson S.J., Bruggeman F.J., Krab K., Schuppen J.H., Hardin H., Bakker B.M., Mone M.J., Rybakova K.N., Eijken M., Leeuwen H.J. and Snoep J.L. (2009) Systems biology towards life in silico: mathematics of the control of living cells. *J. Mathemat. Biol.*, 58, pp. 7-34.
165. Westerhoff H. V., Verma M., Nardelli M., Adamczyk M., van Eunen K., Simeonidis E., Bakker B. M. (2010) Systems biochemistry in practice: experimenting with modelling and understanding, with regulation and control. *Biochemical Society transactions*, 38(5), pp. 1189-96.
166. Wiener N. (1961) Cybernetics; or, control and communication in the animal and the machine, 2nd edn. MIT Press, New York.
167. Wilkinson D.J. (2006) Stochastic Modeling for Systems Biology. Chapman & Hall/CRC, Mathematical and Computational Biology Series, 255 p.
168. Wittig U., Golebiewski, M., Kania, R., Krebs, O., Mir, S., Weidemann, A., Anstein, S., Saric, J. and Rojas, I., (2006) SABIO-RK: Integration and Curation of Reaction Kinetics Data In proceedings of the 3rd International workshop on Data Integration in the Life Sciences 2006 (DILS'06). Hinxton, UK. Lecture Notes in Bioinformatics, 4075: 94-103.
169. Wolkenhauer O., Ullah M., Kolch W., Cho K.H. (2004) Modeling and simulation of intracellular dynamics: Choosing an appropriate framework. *IEEE Tans. Nanobioscience* 3(3), pp. 200-207.
170. Yang, X., Xu, L., Yan, M. (2010) Modeling metabolic response to changes of enzyme amount in yeast glycolysis. *Journal of Biotechnology*, 9(41), pp. 6894-6901.
171. Zerhouni E. (2003) The NIH roadmap. *Science* 302:63-64.