



KAUNO TECHNOLOGIJOS UNIVERSITETAS
ELEKTROS IR ELEKTRONIKOS FAKULTETAS

Dainius Vosylius

DUOMENŲ APROKSIMAVIMO METODŲ TAIKYMAS
BIOTECHNOLOGINIŲ PROCESŲ MODELIAVIMUI

Baigiamasis magistro projektas

Vadovas
Doc. dr. Tomas Tekorius

KAUNAS, 2015

KAUNO TECHNOLOGIJOS UNIVERSITETAS
ELEKTROS IR ELEKTRONIKOS FAKULTETAS
AUTOMATIKOS KATEDRA

DUOMENŲ APROKSIMAVIMO METODŲ TAIKYMAS
BIOTECHNOLOGINIŲ PROCESŲ MODELIAVIMUI

Baigiamasis magistro projektas
M3066L21 Valdymo technologijos (621H66001)

Vadovas

(parašas) Doc. dr. Tomas Tekorius
(data)

Recenzentas

(parašas) Prof. Rimvydas Simutis
(data)

Projektą atliko

(parašas) Dainius Vosylius
(data)

KAUNAS, 2015



KAUNO TECHNOLOGIJOS UNIVERSITETAS

Elektros ir elektronikos fakultetas

(Fakultetas)

Dainius Vosylius

(Studento vardas, pavardė)

M3066L21 Valdymo technologijos (621H66001)

(Studijų programos pavadinimas, kodas)

Baigiamojo projekto „Duomenų aproksimavimo metodų taikymas biotechnologinių procesų modeliavimui“

AKADEMINIO SAŽININGUMO DEKLARACIJA

Kaunas

Patvirtinu, kad mano **Dainiaus Vosylius** baigiamasis projektas tema „Duomenų aproksimavimo metodų taikymas biotechnologinių procesų modeliavimui“ yra parašytas visiškai savarankiškai, o visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs.

Aš suprantu, kad išaiškėjus nesąžiningumo faktui, man bus taikomos nuobaudos, remiantis Kauno technologijos universitete galiojančia tvarka.

(vardą ir pavardę įrašyti ranka)

(parašas)

Vosylius, D. Duomenų aproksimavimo metodų taikymas biotechnologinių procesų modeliavimui. Magistro baigiamasis projektas / vadovas doc. dr. Tomas Tekorius; Kauno technologijos universitetas, elektros ir elektronikos fakultetas, automatikos katedra.

Kaunas, 2015. 45 psl.

SANTRAUKA

Šiame darbe apžvelgiami biotechnologiniai procesai, bioreaktorių tipai, biotechnologinių procesų modeliai ir eksperimentinių duomenų aproksimavimo būdai. Darbe naudojami periodinio su pamaitinimu technologinio proceso eksperimentiniai duomenys. Pagrindinis šio darbo uždavinys iširti aproksimavimo metodų įtaką biotechnologinio proceso modeliavimo tikslumui ir pamaitinimo profilio optimizavimui. Aproksimavimui naudojama tiesinė interpoliacija, interpoliavimas splineu, Ermito interpoliacija, interpoliavimas polinomu ir aproksimavimas logistine funkcija. Matematinio modelio parametrų optimizavimas ir pamaitinimo profilio optimizavimas vykdomas chemotaksis ir evoliucinio metodų deriniu. Aproksimuojant eksperimentinius duomenis logistine funkcija gaunamas mažiausias nuokrypis nuo aproksimuotų duomenų iki modeliuojamų verčių. Optimizuojant pamaitinimo profilį, maksimali tikslinio produkto išeiga buvo didžiausia aproksimuojant polinomu, nors ir modelio nuokrypis nuo aproksimuotų koncentracijos verčių buvo didesnis nei aproksimuojant logistine funkcija.

Reikšminiai žodžiai: aproksimavimas, modeliavimas, chemotaksis, evoliucinis metodas, Ermito, logistinė, polinomas.

Vosylius, Dainius. Application of data approximation methods for biotechnological process modeling. Final project of *master's degree* / supervisor doc. dr. Tomas Tekorius; Kaunas University of Technology, Faculty of Electrical and Electronics Engineering, department of automation.

Kaunas, 2015. 45 psl.

SUMMARY

This Master thesis provides an overview of biotechnological processes, bioreactor types of biotechnological processes, models and experimental data approximation methods. Experimental data is used from periodic batch feeding process. The main objective of this work to investigate the approximation methods on a biotechnological process modeling accuracy and feeding profile optimization. Approximation is done using linear interpolation, spline interpolation, Hermite interpolation, polynomial interpolation and approximation logistic function. Mathematical model and optimization of feeding profile optimization performed chemotaxis and evolutionary methods in combination. Approximation of experimental data with polynomial gives the smallest deviation from the approximated data to modeled values.

Optimizing the feeding profile, the maximum target product yield was achieved with the polynomial approximation, although the deviation from the model approximated concentration values were higher than by logistic function.

Keywords (up to 8 words): approximation , modeling, chemotaxis , evolutionary method, Hermite , logistic , polynomial.

TURINYS

IVADAS	7
1 BIOTECHNOLOGINIŲ PROCESŲ MATEMATINIAI MODELIAI, PARAMETRŲ IDENTIFIKAVIMO METODAI IR DUOMENŲ STATISTINĖ ANALIZĖ	8
1.1 BIOTECHNOLOGINIAI PROCESAI	8
1.2 BIOREAKTORIAI	8
1.2.1 Anaerobiniai ir aerobiniai bioreaktoriai	8
1.2.2 Nenutrūkstamo veikimo bioreaktoriai (angl. continuous)	9
1.2.3 Periodinio veikimo bioreaktoriai (angl. batch)	9
1.2.4 Pusiaus periodinio veikimo sistemos (angl. semi-batch)	9
1.2.5 Reaktoriai su suspenduota mikroorganizmų mase (angl. suspended cell reactor)	9
1.2.6 Bioreaktoriai su imobilizuota mikroorganizmų mase (angl. immobilized cell reactor)	9
1.3 BIOTECHNOLOGINIO PROCESO MODELIAVIMAS.	10
1.3.1 Medžiagų masės balansas	10
1.3.2 Kinetiniai (mechanistiniai) modeliai	10
1.3.3 Biomasės išėigos koeficientas ir gyvybinių funkcijų palaikymas	11
1.3.4 Negyvos ląstelės augančioje mikroorganizmų kultūroje	11
1.3.5 Inhibavimas	12
1.4 DARBE NAUDOJAMAS MODELIS.	12
2 BIOTECHNOLOGINIŲ PROCESŲ MATEMATINIŲ MODELIŲ PARAMETRŲ IDENTIFIKAVIMO METODAI	13
2.1 APROKSIMAVIMAS.....	13
1.1.1. Aproksimavimas interpoliuojant polinomu.	14
2.1.1 Aproksimavimas kubine Hermito interpoliacija.	18
2.1.2 Aproksimavimas splineu.....	20
2.1.3 Aproksimavimas logistine funkcija	22
2.2 PARAMETRŲ IDENTIFIKAVIMAS PIRMU PRIARTĖJIMU	23
2.3 PARAMETRŲ IDENTIFIKAVIMAS.	26
2.4 PAMAITINIMO PROFILIO OPTIMIZAVIMAS.....	31
3 BIOTECHNOLOGINIŲ PROCESŲ MATEMATINIAI MODELIAVIMO REZULTATAI MATLAB APLINKOJE	32
3.1 PARAMETRŲ VERTĖS PIRMU PRIARTĖJIMU:	32
3.2 PARAMETRŲ OPTIMIZAVIMAS CHEMOTAKSIS IR EVOLIUCINIO METODŲ DERINIU (MUTACIJŲ SKAIČIUS 10)	33
3.3 PAMAITINIMO PROFILIO OPTIMIZAVIMAS:.....	38
4 IŠVADOS	44
5 LITERATŪRA	45

Ivadas

Prieš parenkant matematinio modelio struktūrą, identifikuojant matematinio modelio parametrų vertes reikia atlikti modeliuojamų eksperimentinių duomenų analizę. Praktikoje pamatuoti biotechnologinių procesų santykinų reakcijos greičių neįmanoma, jie yra išskaičiuojami iš tam tikrų išraiškų. Biotechnologinio proceso išsaugomi eksperimentiniai duomenys dažniausiai būna medžiagų koncentracijos, kultivavimo terpės tūris ir tiesiogiai matuojami fizikiniai ir cheminiai duomenys. Eksperimentiniai duomenys, tokie kaip koncentracija, matuojami tam tikrais laiko intervalais, tačiau matematinio modelio parametrų identifikavimui reikalingos ir eksperimentinių duomenų vertės tarp matavimų. Šios vertės yra gaunamos aproksimuojant eksperimentinius duomenis, tačiau rezultatas gali būti labai skirtingas priklausomai nuo aproksimavimo būdo. Todėl modelio tikslumas, o taip pat ir maksimali tikslinio produkto išeiga priklauso nuo eksperimentinių duomenų aproksimavimo būdo.

Tyrimų objektas:

Duomenų aproksimavimo metodų įtaka biotechnologinio proceso modeliavimui.

Darbo tikslas:

Tirti eksperimento duomenų aproksimavimo metodų taikymą biotechnologinių procesų modeliavimui.

Uždaviniai:

1. Apžvelgti bioreaktorių, biotechnologinių procesų matematinius modelius, jų parametrų identifikavimo metodus ir išanalizuoti eksperimento duomenų aproksimavimo metodus.
2. Suderinti programą MATLAB aplinkoje, kuri keliais metodais aproksimuotų eksperimento duomenis, identifiкуotų biotechnologinio proceso modelio parametrus, optimizuotų pamaitinimo profilį.
3. Nustatyti aproksimavimo metodų įtaką modeliavimo tikslumui ir pamaitinimo profilio optimizavimui, panaudojant sukurtą programą ir realaus biotechnologinio proceso eksperimento duomenis.

1 Biotechnologinių procesų matematiniai modeliai, parametrų identifikavimo metodai ir duomenų statistinė analizė.

1.1 Biotechnologiniai procesai.

Biotechnologija - tai mokslo šaka, nagrinėjanti gyvųjų organizmų ir biologinių procesų taikymo technologijoje problemas. Biotechnologijos mokslas susijęs su produktų gavimu iš gyvųjų organizmų ir biologinių procesų. Žmonės su biotechnologiniais procesais susidūrė labai senai, dar prieš atsirandant biotechnologijos terminui. Pirmieji žinomi biotechnologiniai procesai buvo sūrių, vyno ir alaus gamyba. Biotechnologinių procesų taikymas praktikoje plačiai išplito apie XX a.

Šie procesai yra vieni sudėtingiausi technologinių procesų. Juose vykstant šimtams biocheminių reakcijų, daroma įtaka gamybos sąnaudoms, galutiniam produkto kiekiui ir kokybei. Esant didelei ekonominei biotechnologinių procesų svarbai, būtina užtikrinti optimalų juose vykstančių reakcijų darbą, maksimizuojant kokybės ir ekonominius rodiklius, nes lyginant cheminių reakcijų ir biotechnologinių procesų būdu gauto produkto sąnaudų ir pelno santykį pastarasis yra geresnis.

Sparčiai tobulėjant informacinėms ir valdymo technologijoms atsiveria galimybės maksimaliai tiksliai valdyti ir optimizuoti sudėtingiausius biotechnologinius procesus, kurie taikomi keturiose pagrindinėse srityse: medicinoje, augalininkystėje ir žemės ūkyje, augalininkystėje ne maisto reikmėms (biokuro, yrančio plastiko, augalinei alyvai ir pan.) ir aplinkosaugai.

1.2 Bioreaktoriai.

[1] Bioreaktorius – tai indas, kuriame vyksta biocheminės reakcijos su mikroorganizmais arba fermentais.

Pagal požymius, naudojamas vykstančiuose procesuose deguonis ar ne, jo tiekimo ir maišymo ypatybės, biomasės būsenas, iš dalies grąžinama atgal į bioreaktorių ar imobilizuojama, žaliavos tiekimo į bioreaktorių režimą, bioreaktoriai skirstomi į:

1.2.1 Anaerobiniai ir aerobiniai bioreaktoriai.

Anaerobiniuose reaktoriuose vykstantys biotechnologiniai procesai reikalauja, kad nepatektų deguonis.

Aerobiniuose reaktoriuose biomasė specialiai yra aprūpinama deguonimi [1].

1.2.2 Nenutrūkstamo veikimo bioreaktoriai (angl. continuous).

Tai reaktorius į kurį įtekančių skysčių srauto suma yra lygi ištekančios terpės srauto sumai. Šio tipo reaktoriai skirstomi į idealaus sumaišymo ir tiesinius reaktorius [1].

1.2.3 Periodinio veikimo bioreaktoriai (angl. batch).

Periodinio veikimo bioreaktoriuose kultivavimo procesas vykdomas periodiškai. Pvz. procesas susideda iš reaktoriaus sterilizavimo, užsėjimo, kultivavimo proceso, pagaminto produkto pašalinimo ir reaktoriaus paruošimo kitam ciklui.

Šio tipo bioreaktoriuose gali būti vykdomi periodiniai arba periodiniai su pamaitinimu kultivavimo procesai. Pirmuoju atveju pamaitinimo substratai į kultivavimo terpę prieš užsėjant mikroorganizmų kultūrą, o proceso metu pamaitinimas nevykdomas. Antruoju atveju pamaitinimo substratų srautai tiekiami proceso metu. Dėl šių proceso vykdymo ypatumų šiuose bioreaktoriuose gali būti pasiektos didesnės biomasės ar tikslinio produkto koncentracijos (ypač periodinio su pamaitinimu proceso metu), lyginant su atitinkamos kultūros kultivavimo procesu nenutrūkstamo veikimo būdu. Dėl to palengvėja tikslinio produkto gryninimo procesas [1].

1.2.4 Pusiau periodinio veikimo sistemos (angl. semi-batch).

Pilnai užpildytame bioreaktoriuje kultivavimo procesas vykdomas tol kol gaunamas norimas biomasės kiekis ar tikslinio produkto koncentracija. Iš bioreaktoriaus pašalinama didžioji dalis kultivavimo terpės ir dalis terpės paliekama reaktoriuje sekančiam kultivavimo ciklui įvykdyti. Reaktorius vėl užpildomas reikiama substratais ir kitom būtinom medžiagom kultivavimo procesui vykdyti [1].

1.2.5 Reaktoriai su suspenduota mikroorganizmų mase (angl. suspended cell reactor).

Biomasė suspenduojama kultivavimo terpėje ir vykdant procesą biomasė iš dalies gražinama atgal į bioreaktorių. Šis būdas taikomas biologinio nuotekų valymo procesuose [1].

1.2.6 Bioreaktoriai su imobilizuota mikroorganizmų mase (angl. immobilized cell reactor).

Bioreaktoriuje mikroorganizmai yra įtvirtinami ant įvairių paviršių ir pašalinant terpę iš reaktoriaus, mikroorganizmai lieka kultivavimo terpėje. Šis būdas naudojamas aerobiniuose ir pratekančio skysčio bioreaktoriuose [1].

1.3 Biotechnologinio proceso modeliavimas.

1.3.1 Medžiagų masės balansas.

Atliekant biotechnologinių procesų modeliavimą daugeliu atvejų reikalinga apsirašyti medžiagų masės balanso ir energijos srautų balanso lygtis.

Medžiagos masės lygtys paremtos medžiagų tvermės dėsniu, kuris šiuo atveju reiškia, kad tam tikros medžiagos kiekio kitimas reaktoriuje gali kisti dėl:

- Reaktoriuje vykstančių cheminių ar biocheminių reakcijų, kuomet medžiaga skyla arba dalijasi;
- Dėl išorinio medžiagų tiekimo į bioreaktorių;
- Iš bioreaktoriaus pašalinamo kultivavimo terpės kiekio.

$$\frac{d(C_1V)}{dt} = q_1C_2V + C_{in1}F_{in} - C_1F_{out} \quad (1.1)$$

Kairioji lygybės pusė nusako C_1 medžiagos kiekio kitimą laike [g/h]. q_1 – santykinis reakcijos greitis medžiagos C_2 atžvilgiu [g C_1 /g C_2 h]. C_{in1} – įtekančios medžiagos koncentracija (pvz. pamaitinant) [g/l]. F_{in} – įtekančios medžiagos kiekio srautas [l/h]. C_1 – bioreaktoriuje esančios medžiagos koncentracija prieš pašalinant (pvz. imant mėginius) [g/l]. F_{out} – pašalinamos medžiagos kiekio srautas [l/h] [1].

1.3.2 Kinetiniai (mechanistiniai) modeliai.

[1] Dažniausiai naudojama Mono kinetika nusako santykinio biomasės augimo greičio μ priklausomybę nuo substrato koncentracijos s .

Monod;

$$\mu = \mu_{max} \frac{s}{K_s + s} \quad (1.2)$$

μ - santykinis biomasės augimo greitis;

μ_{max} – parametras parenkamas pagal konkretų mikroorganizmą (pvz. E.coli $\mu_{max} = 0,5$ [l/h]);

K_s - parametras parenkamas pagal konkretų mikroorganizmą (pvz. E.coli $K_s = 4$ [mg/L]);

Literatūroje yra daugiau kinetinių modelių. Dažniausiai naudojamos alternatyvių kinetinių modelių priklausomybės:

Blackman:

$$\mu = \mu_{max} \min\left(1, \frac{s}{2K_s}\right) \quad (1.3)$$

Teissier:

$$\mu = \mu_{max} \left(1 - \exp\left(-\frac{s}{K_T}\right) \right) \quad (1.4)$$

Moser:

$$\mu = \mu_{max} \frac{s^n}{K_S + s^n} \quad (1.5)$$

Contois:

$$\mu = \mu_{max} \frac{s}{xK_C + s} \quad (1.6)$$

Powel:

$$\mu = \mu_{max} \frac{s - K_1 t(s)}{K_S + s - K_1 t(s)} \quad (1.7)$$

1.3.3 Biomassės išėigos koeficientas ir gyvybinių funkcijų palaikymas.

Biomassė auga dėl sunaudojamo substrato, todėl santykinis substrato suvartojimo greitis yra proporcingas santykiniai biomassės augimo greičiui:

$$q = \frac{\mu}{Y_{xs}} \quad (1.8)$$

Tačiau egzistuoja dar substrato suvartojimas m gyvybinių funkcijų palaikymui:

$$q = \frac{\mu}{Y_{xs}} + m \quad (1.9)$$

Substrato suvartojimas gyvybinių funkcijų palaikymui yra ypač aktualus, kai biomassės santykinis augimo greitis sumažėja ir didžioji dalis substrato suvartojama gyvybinėms funkcijoms palaikyti [1].

1.3.4 Negyvos ląstelės augančioje mikroorganizmų kultūroje.

Augant mikroorganizmų kultūrai dalis ląstelių nebenaudoja substratų ir nustoja augti, kitaip tariant jos tampa negyvos. Negyvų ląstelių atsiradimą lemia įvairūs fizikiniai ir cheminiai poveikiai, kaip per aukšta temperatūra, netinkamas pH lygis, nuodingų medžiagų poveikis.

Daroma prielaida, jog ląstelių žuvimas yra proporcingas gyvų ląstelių koncentracijai [1]:

$$\frac{dx}{dt} = -k_d x \quad (1.10)$$

k_d - konstanta [1/h]

1.3.5 Inhibavimas.

Inhibavimas, kitaip tariant slopinimas – mažina santykinius biomasės augimo ir substrato sunaudojimo greičius. Inhibavimas dažniausiai pasireiškia dėl šalutinio produkto. Inhibavimas gali būti vertinamas naudojant tam tikrą daugiklį (pvz. Ierusaliminskio, Hinshelwoodo, Levenspilio, Aibos) arba naudojant Pirtto ar Bellgardto pasiūlytas išraiškas. Mes naudosime Bellgardto pasiūlytą išraišką [1]:

$$q = q_{max} \frac{s}{K_s + s + \frac{s^2}{K_{si}}} \quad (1.11)$$

Substrato koncentracija, užtikrinanti maksimalų santykinį substrato sunaudojimo greitį paskaičiuojama:

$$s_{max} = \sqrt{K_{si}K_s} \quad (1.12)$$

Maksimalus santykinis substrato suvartojimo greitis paskaičiuojamas pagal:

$$\frac{q}{q_{max}} = \frac{\sqrt{K_{si}}}{\sqrt{K_{si} + 2\sqrt{K_s}}} \quad (1.13)$$

1.4 Darbe naudojamas modelis.

Dinaminio modelio struktūra pasirinkta vadovaujantis kinetine sąveika, kuri naudojama mechanistiniuose modeliuose biomasės santykinio augimo ir substrato santykinio suvartojimo vienetais nuspėti. Pasirinktas modelis yra sudarytas iš kelių tipinių mechanistinių modelių, kurie dažniausiai naudojami periodinio veikimo ir periodinio veikimo su pamaitinimu bioreaktoriuose [2] [3];

Biomasės masės balanso lygtis:

$$\frac{dx}{dt} = \mu(s, x)x - u \frac{x}{V}, \quad (1.14)$$

Substrato masės balanso lygtis:

$$\frac{ds}{dt} = -q(s, x)x + u \frac{s_f - s}{V}, \quad (1.15)$$

Kultivavimo terpės tūrio balanso lygtis:

$$\frac{dV}{dt} = u. \quad (1.16)$$

Biomasės santykinis augimo greitis įvertinant mirusiųjų ląstelių santykinį didėjimo greitį:

$$\mu(s, x) = \mu_g(s, x) - \mu_d(x) \quad (1.17)$$

Biomasės santykinis augimo greitis nevertinant mirusiųjų ląstelių:

$$\mu_g(s, x) = \frac{\mu_{max} \cdot s}{\left(k_s + s + \frac{s^2}{k_i}\right)(1 + \kappa_1 x)} \quad (1.18)$$

Mirusiųjų ląstelių santykinis didėjimo greitis:

$$\mu_d(x) = (k_x + \kappa_2 x) \quad (1.19)$$

Substrato santykinis suvartojimo greitis:

$$q(s, x) = (Y_{xs} + \kappa_3 x)\mu_g(s, x) + \kappa_4 \quad (1.20)$$

x – biomasės koncentracija;

s – substrato koncentracija;

μ - biomasės santykinis augimo greitis;

μ_g – biomasės santykinis augimo greitis neįvertinant mirusiųjų ląstelių;

μ_d – mirusiųjų ląstelių santykinis didėjimo greitis;

q – substrato santykinis suvartojimo greitis;

$k_s, k_i, k_x, Y_{xs}, \kappa_1, \kappa_2, \kappa_3, \kappa_4$ ir μ_{max} yra modelio parametrai, kuriuos reikia identifikuoti.

2 Biotechnologinių procesų matematinių modelių parametrų identifikavimo metodai [4].

Parametrai kuriuos būtina identifikuoti yra $k_s, k_i, k_x, Y_{xs}, \kappa_1, \kappa_2, \kappa_3, \kappa_4$. Parametrų identifikavimas susideda iš kelių etapų.

2.1 Aproksimavimas.

Pirmiausia reikia aproksimuoti turimus eksperimentinius duomenis. Funkcijų aproksimavimo uždavinių gausu įvairiose matematikos, fizikos ir technikos srityse. Literatūroje nagrinėjama daugybė funkcijų aproksimavimo uždavinio sprendimo metodų. Tai paaiškinama tuo, kad praktikoje susiduriama su daugeliu skirtingų šio uždavinio formuluočių.

Apibendrintai funkcijų aproksimavimo uždavinį galima suformuluoti taip:

turint funkcijos $y = f(x)$ reikšmių lentelę (x_i, y_i) (čia $i = 0, n, y_i = f(x_j)$ ir $x_i \neq x_j$, kai $i \neq j$), reikia rasti funkcijos $f(x)$ reikšmę, kai $x = t$, o $t \in [x_0; x_n]$.

Funkcijos $f(x)$ analizinė išraiška paprastai yra nežinoma arba pernelyg sudėtinga.

Atsižvelgiant į y_i reikšmių tikslumą, taikomi du šio uždavinio sprendimo metodai:

- interpoliavimo metodas;

- suglodinimo metodas.

Interpoliavimas tai aproksimuojančios funkcijos radimas, kuri kerta visus eksperimentinius taškus.

Suglodinimas naudojamas, kai manoma, jog eksperimentiniai duomenys turi paklaidas, tuomet naudoti funkcija kuri kerta visus eksperimentinius taškus netikslinga, geriau rasti tokią funkciją, kuri pagal pasirinktą kriterijų geriausiai aproksimuotų.

Plačiau bus apžvelgta duomenų aproksimavimas tiesiškai interpoliuojant, aproksimavimas polinomu (daugianariais), Ermito interpoliacija, interpoliavimas splineu, aproksimavimas logistine funkcija. Be šių aproksimavimo metodų yra tikrai daug kitų, tai Gauso, Oilerio, Esscherto, Teiloro, Liovel'o, Dirichlet'o ir Kronsker'io, Diofatinė aproksimacijos, aproksimavimas apskritimais, laiptinėmis funkcijomis (reguliariųjų, be antros rūšies trūkių funkcijų) bei mažiausių kvadratų metodai.

1.1.1. Aproksimavimas interpoliuojant polinomu.

Polinomas (poli – daug, nomen – vardas, pavadinimas) kitaip tariant daugianaris yra sudarytas iš kelių narių pvz. $a_0 + a_1x$. Interpoliavimas, tai yra tarpinių taškų radimas tarp jau turimų taškų. Polinominis interpoliavimas tai yra tinkamiausios polinominės tiesės ar kreivės, kuri atitiktų turimus taškus, radimas [5]. Polinominis interpoliavimas gali būti kelių laipsnių, pvz.:

Pirmo laipsnio polinomas (tiesinė interpoliacija):

$$f(x) = a_0 + a_1x \quad (2.1)$$

Koeficientams a_0, a_1 surasti reikia dviejų turimų taškų. Tai yra tiesės funkcija, kuri jungia du ar daugiau taškų, esančių toje pačioje tiesėje.

Antro laipsnio polinomas (kvadratinė interpoliacija):

$$f(x) = a_0 + a_1x + a_2x^2 \quad (2.2)$$

Koeficientams a_0, a_1, a_2 surasti reikia trijų turimų taškų.

Trečio laipsnio polinomas (kubinė interpoliacija):

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \quad (2.3)$$

Koeficientams a_0, a_1, a_2, a_3 surasti reikia keturių turimų taškų.

Kuo daugiau eksperimentinių taškų bandysime interpoliuoti polinomu, tuo sudėtingesnė ir aukštesnio laipsnio bus kreivės funkcija.

Vienas iš paprasčiausių būdų interpoliuoti yra tiesinis interpoliavimas, nes tam reikia turėti tik du žinomus taškus. Tarpiniai taškai tarp dviejų turimų taškų yra randami iš funkcijos:

$$f(x) = a_0 + a_1x \quad (2.4)$$

arba mūsų atveju:

$$x(t) = a_0 + a_1 t \quad (2.5)$$

Koeficientai a_0 ir a_1 yra randami iš lygčių sistemos, kuri sudaroma iš dviejų turimų taškų:

Turime biomasės augimo ir substrato suvartojimo eksperimentinius duomenis:

2.1 lentelė. Biomasės koncentracijos eksperimentiniai duomenys.

x(t)	0,104	0,725	4,783	10,000	16,377	27,391	47,971	74,638	96,667	129,420	123,333
t	0	2,938	4,900	7,000	7,969	8,931	9,932	10,896	11,904	12,865	13,870

2.2 lentelė. Substrato koncentracijos eksperimentiniai duomenys.

s(t)	8,870	8,581	7,847	5,569	1,670	3,948	6,129	7,828	4,517	0,531	0,0001
t	0	2,938	4,900	7,000	7,969	8,931	9,932	10,896	11,904	12,865	13,870

Taigi sudarę ir išsprendę lygčių sistemą iš dviejų turimų eksperimentinių taškų (2,938;0,725) ir (4,900; 4,783), rasime ieškomus koeficientus:

$$x(2,938) = 0,725 = a_0 + a_1 \cdot 2,938 \quad (2.6)$$

$$x(4,900) = 4,783 = a_0 + a_1 \cdot 4,900 \quad (2.7)$$

$$\begin{bmatrix} 1 & 2,938 \\ 1 & 4,900 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0,725 \\ 4,783 \end{bmatrix} \quad (2.8)$$

Išsprendę lygčių sistemą randame a_0 ir a_1 koeficientų vertes:

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 1,106 \\ -0,130 \end{bmatrix} \quad (2.9)$$

Jas įstatę į pradinę funkciją gauname tiesės funkciją, kuri jungia du taškus (2,938;0,725) ir (4,900; 4,783):

$$x(t) = 1,106 - 0,103 \cdot t \quad (2.10)$$

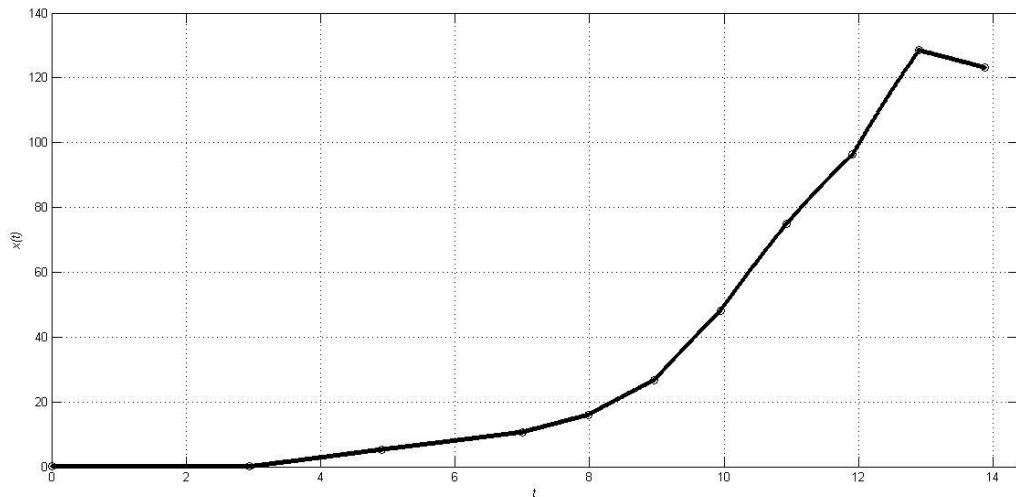
Jei biotechnologinis procesas kinta netiesiškai, randamos tiesinės funkcijos tarp kiekvienos taškų poros.

Naudojantis MATLAB programa nubrėžėme biomasės koncentracijos augimo ir substrato koncentracijos mažėjimo kreives tiesinio interpoliavimo metodu.

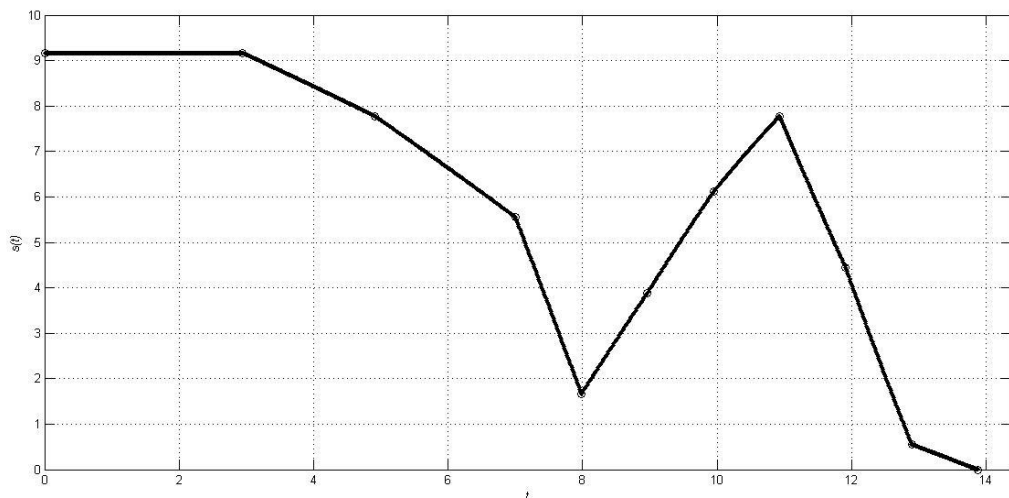
```
% x(t) - biomasės koncentracija
% s(t) - substrato koncentracija
t=[0 2.938 4.900 7.000 7.969 8.931 9.932 10.896 11.904 12.865 13.870];
x=[0.104 0.725 4.783 10.000 16.377 27.391 47.971 74.638 96.667 129.420
123.333];
s=[8.870 8.581 7.847 5.569 1.670 3.948 6.129 7.828 4.517 0.531 0.0001];

step=0.001; %žingsnis
```

```
tt=0:step:14; %intervalas
x_grafikas=interp1(t, x, tt, 'linear'); %biomasės augimo interpoliavimas
s_grafikas=interp1(t, s, tt, 'linear'); %substrato mažėjimo
interpoliavimas
plot(t, x, 'o', tt, x_grafikas); grid on; %biomasės grafiko braižymas
figure
plot(t, s, 'o', tt, s_grafikas); grid on; %substrato grafiko braižymas
```



2.1 pav. Biomasės koncentracijos kitimo grafikas laiko atžvilgiu tiesiškai interpoliuotas.



2.2 pav. Substrato koncentracijos kitimo grafikas laiko atžvilgiu tiesiškai interpoliuotas.

Tiesinis interpoliavimo metodas labiau tinka procesams, kurie kinta tiesiškai. Kadangi mūsų grafikas nėra tiesinis, ši metodika nebus naudojama.

Aukštesnio laipsnio polinominis interpoliavimas naudojamas aproksimuoti sudėtingoms kreivėms. Norint rasti antros eilės polinomą, reikia rasti tris koeficientus a_0, a_1, a_2 :

$$f(x) = a_0 + a_1x + a_2x^2 \quad (2.11)$$

a_0, a_1, a_2 koeficientų radimui yra sudaroma ir išsprendžiama trijų lygčių sistema iš trijų eksperimentinių taškų, pavyzdžiui (2,938; 0,725), (4,900; 4,783) ir (7,000; 10,000):

$$x(2,938) = 0,725 = a_0 + a_1 \cdot 2,938 + a_2 \cdot 2,938^2 \quad (2.12)$$

$$x(4,900) = 4,783 = a_0 + a_1 \cdot 4,900 + a_2 \cdot 4,900^2 \quad (2.13)$$

$$x(7,000) = 10,000 = a_0 + a_1 \cdot 7,000 + a_2 \cdot 7,000^2 \quad (2.14)$$

$$\begin{bmatrix} 1 & 2,938 & 8,632 \\ 1 & 4,900 & 24,010 \\ 1 & 7,000 & 49,000 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0,725 \\ 4,783 \\ 10,000 \end{bmatrix} \quad (2.15)$$

Išsprendę lygčių sistemą ir nustatę koeficientų a_0, a_1, a_2 vertes, gauname kvadratinę funkciją kuri jungia tris mūsų eksperimento taškus:

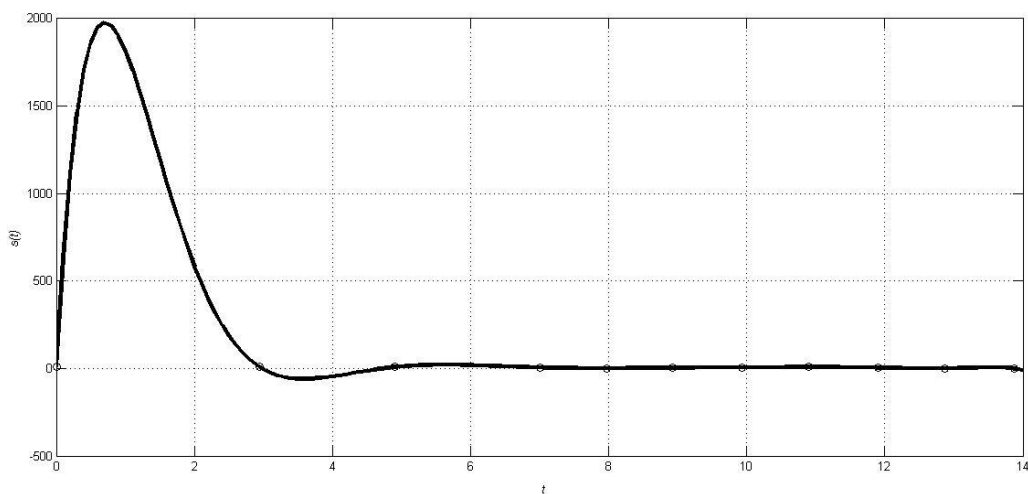
$$x(t) = -3,877 + 1,266x + 0,102t^2 \quad (2.16)$$

Analogiškai gali būti randami aukštesnių polinomų funkcijų koeficientai.

Interpoliaciniam daugianariui (polinomui) apskaičiuoti kiekvieną kartą spręsti aukšto laipsnio lygčių sistemą būtų neracionalu. Daugianariams paskaičiuoti Lagranžas yra pasiūlęs išraišką:

$$L(x) = \sum_{j=0}^k y_j \prod_{\substack{0 \leq m \leq k \\ m \neq j}} \frac{x-x_m}{x_j-x_m} \quad (2.17)$$

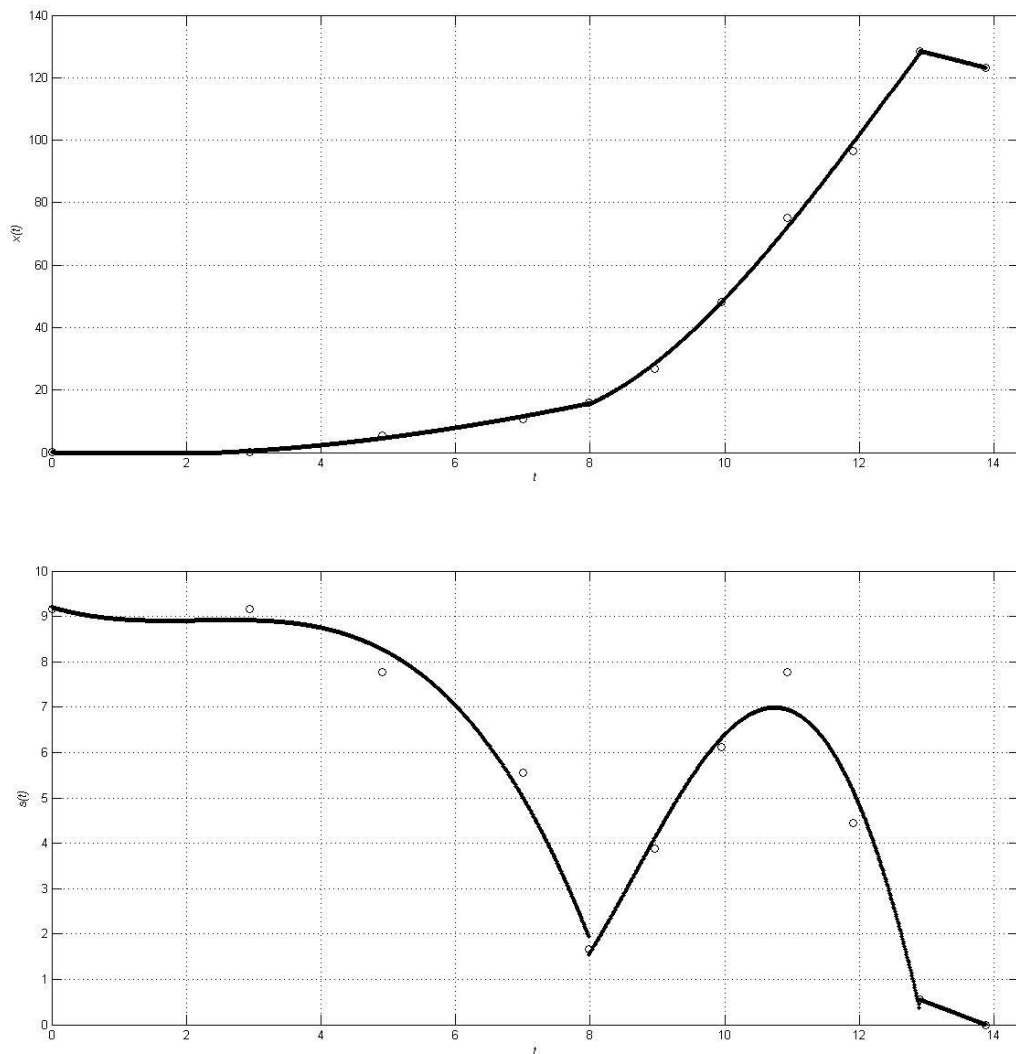
Naudojantis MATLAB programa aproksimuosime substrato koncentracijos eksperimentinius duomenis 10 eilės polinomu.



2.3 pav. biomasės koncentracijos kitimo grafikas laiko atžvilgiu interpoliuojant 10 eilės polinomu.

Kaip matome iš grafiko koncentracijos vertė pakyla netoli 2000. Taip įvyko todėl, kad matavimo duomenys yra pamatuoti ne vienodais laiko tarpais. Toks grafiko iškreipimas interpoliuojant polinomu vadinamas Rungės fenomenu [6].

Rungės fenomenui išvengti aproksimuosime dalimis 3 laipsnio polinomu. Biomasės koncentracijos kitimo grafikas aproksimuojamas iš dviejų dalių, nuo 0 iki 12 valandos trečiu laipsniu, o nuo 12 iki 13,870 valandos aproksimuojama pirmu laipsniu. Substrato grafikas aproksimuojamas iš trijų dalių – nuo 0 iki 8 valandų ir nuo 8 iki 12,865 trečiu laipsniu, o nuo 12,865 iki 13,870 valandos aproksimuojama pirmu laipsniu:



2.1.1 Aproksimavimas kubine Hermito interpoliacija.

Interpoliuojant Hermito daugianariu reikia rasti daugianarį, kuris eina ne tik per duotus taškus, bet ir tam tikras išvestines tuose taškuose [7]. Jei be $f(x_i)$ verčių mazge x_i žinomos ir $f'(x)$

„Duomenų aproksimavimo metodų taikymas biotechnologinių procesų modeliavimui“

išvestinių iki $(m_i - 1)$ -osios eilės imtinai vertės, tai galima reikalauti, kad interpoliacinis daugianaris $y = F(x)$ tenkintų Hermito interpoliacijos sąlygas:

$$F^{(l)}(x_i) = f^{(l)}(x_i), \quad i = \overline{0, n}, l = \overline{0, m_i - 1}$$

Simbolis l žymi l -osios eilės išvestinę. Daugianaris tenkinantis šias sąlygas, vadinamas Hermito interpoliaciniu daugianariu. Jo laipsnis turi būti

$$N = m_0 + m_1 + \dots + m_n - 1$$

Jo koeficientai apskaičiuojami iš lygčių sistemos, kurią apibrėžia Hermito interpoliacijos sąlygos.

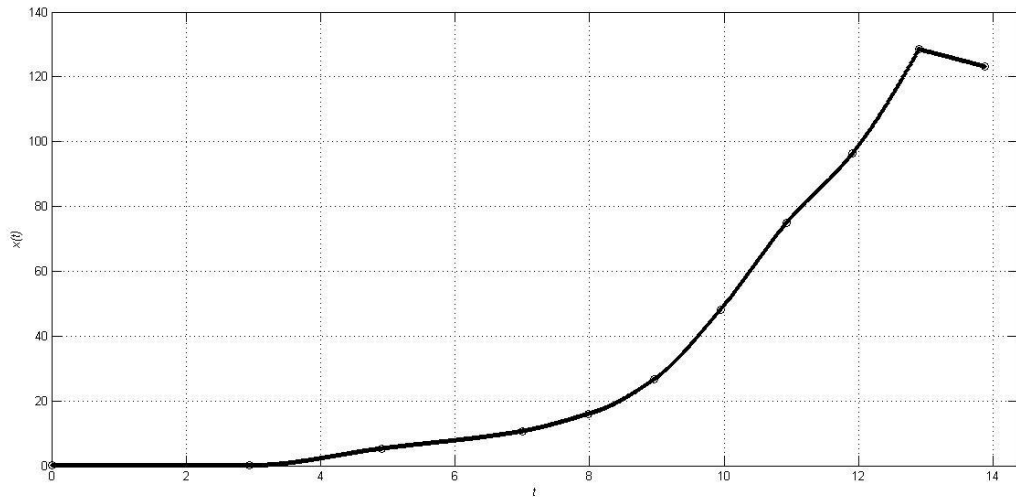
Hermito interpoliacinio daugianario išraiškos gali būti sudaromos Lagranžo arba Niutono daugianariais [8].

Matlab aplinkoje Hermito interpoliacimas vykdomas dalimis 3 laipsniu (pchip – piecewise cubic Hermite interpolation) [9].

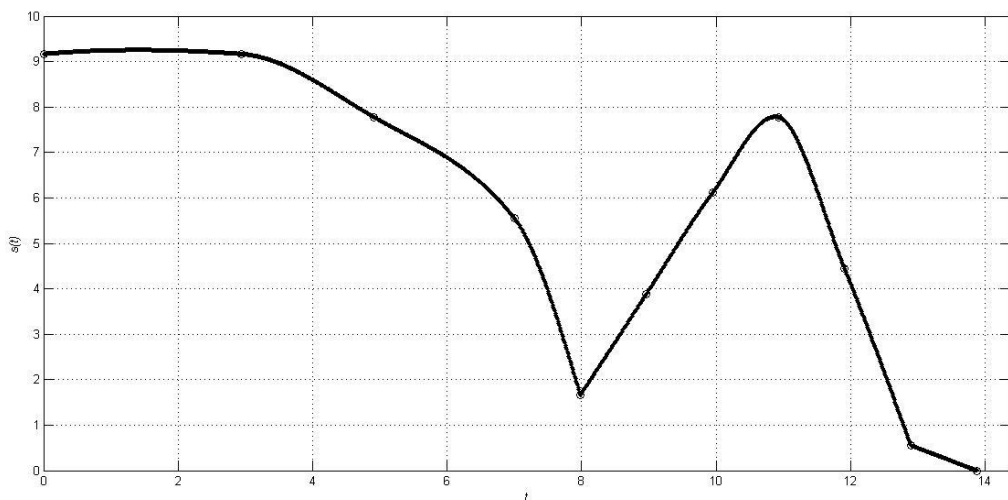
Duomenų aproksimavimas Hermito interpoliacimu MATLAB aplinkoje aprašomas:

```
% x(t) - biomases koncentracija
% s(t) - substrato koncentracija
t=[0 2.938 4.900 7.000 7.969 8.931 9.932 10.896 11.904 12.865 13.870];
x=[0.104 0.725 4.783 10.000 16.377 27.391 47.971 74.638 96.667 129.420
123.333];
s=[8.870 8.581 7.847 5.569 1.670 3.948 6.129 7.828 4.517 0.531 0.0001];

step=0.001; %žingsnis
tt=0:step:14; %intervalas
x_grafikas=interp1(t, x, tt, 'pchip'); %biomasės augimo interpoliacimas
s_grafikas=interp1(t, s, tt, 'pchip'); %substrato mažėjimo
interpoliacimas
plot(t, x, 'o', tt, x_grafikas); grid on; %biomasės grafiko braižymas
figure
plot(t, s, 'o', tt, s_grafikas); grid on; %substrato grafiko braižymas
```



2.4 pav. biomasės koncentracijos kitimo grafikas laiko atžvilgiu interpoliuotas Hermito daugianariais.



2.5 pav. substrato koncentracijos kitimo grafikas laiko atžvilgiu interpoliuotas Hermito daugianariais.

2.1.2 Aproksimavimas splineu.

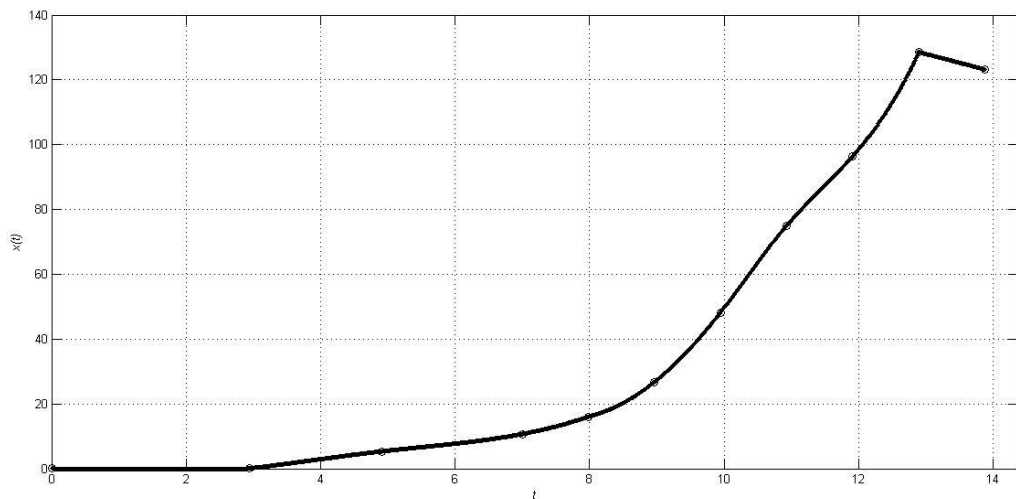
Splinas naudoja žemo laipsnio polinomus kiekviename intervale ir pasirenka tinkamiausias polinomo dalis, kurios tarpusavy puikiai atitinka. Tokiu būdu yra išvengiama Rungės fenomeno [10].

Grafiką nubrėšime naudojantis MATLAB programa. Ketvirtąjį funkcijos `interp1` argumentą pakeičiame į `'spline'`:

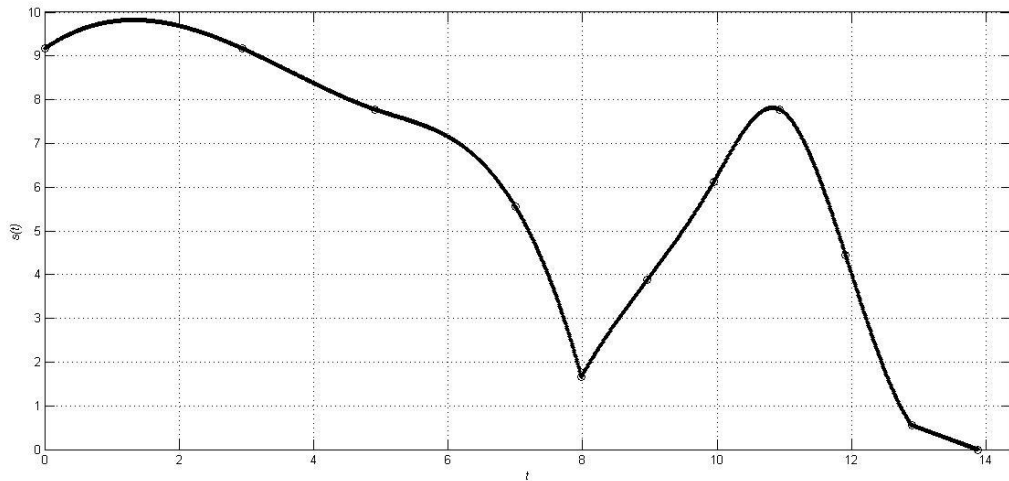
„Duomenų aproksimavimo metodų taikymas biotechnologinių procesų modeliavimui“

```
% x(t) - biomasės koncentracija
% s(t) - substrato koncentracija
t=[0 2.938 4.900 7.000 7.969 8.931 9.932 10.896 11.904 12.865 13.870];
x=[0.104 0.725 4.783 10.000 16.377 27.391 47.971 74.638 96.667 129.420
123.333];
s=[8.870 8.581 7.847 5.569 1.670 3.948 6.129 7.828 4.517 0.531 0.0001];

step=0.001; %žingsnis
tt=0:step:14; %intervalas
x_grafikas=interp1(t, x, tt, 'spline'); %biomasės augimo interpoliavimas
s_grafikas=interp1(t, s, tt, 'spline'); %substrato mažėjimo
interpoliavimas
plot(t, x, 'o', tt, x_grafikas); grid on; %biomasės grafiko braižymas
figure
plot(t, s, 'o', tt, s_grafikas); grid on; %substrato grafiko braižymas
```



2.6 pav. biomasės koncentracijos kitimo grafikas laiko atžvilgiu aproksimuotas splineu.



2.7 pav. substrato koncentracijos kitimo grafikas laiko atžvilgiu aproksimuotas splainu.

Interpoliuoto grafiko polinomu ir splainu tarpiniai taškai yra tikslesni lyginant su tiesiškai interpoliuotu grafiku. Taip pat splainu skaičiavimas yra paprastesnis nei interpoliavimas polinomu, nes nereikia skaičiuoti aukšto laipsnio polinomais.

2.1.3 2.1.4 Aproksimavimas logistine funkcija [11].

Logistinė funkcija yra S formos (riestinė kreivė (angl. sigmoidal curve)) kreivė.

$$y = \frac{b}{[1 + \exp(\sum_{l=0}^m a_l t^l)]} \quad (2.18)$$

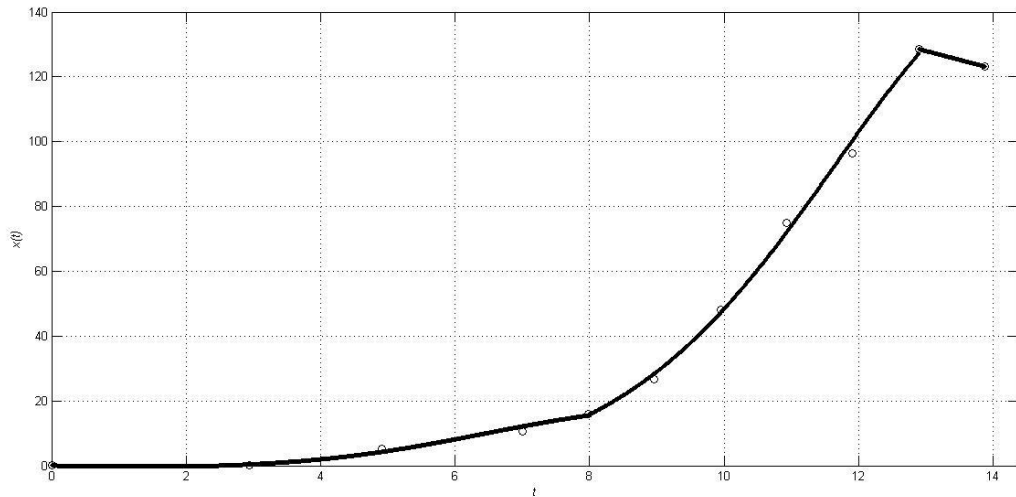
t – laikas, h;

a, b – ieškomi funkcijos parametrai;

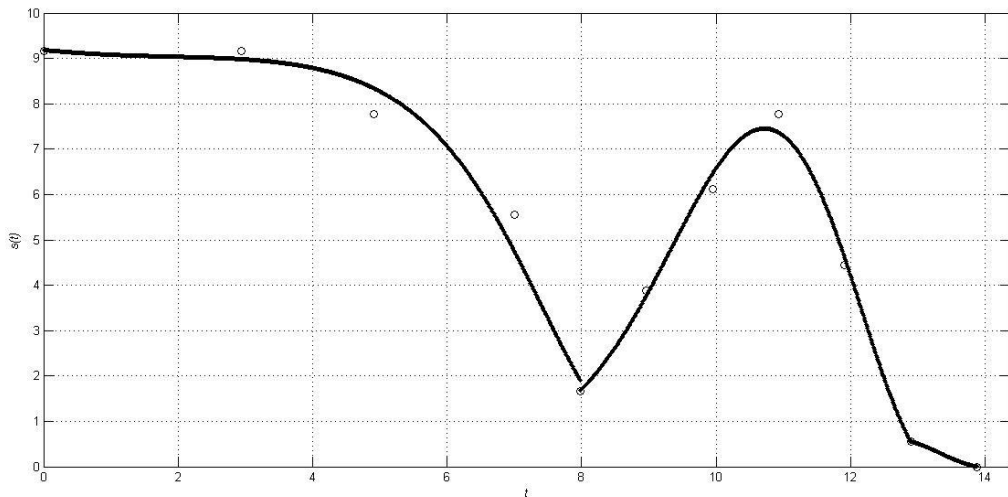
m – daugianario laipsnis.

Išraiškos vardiklyje \exp funkcijos argumentas yra daugianaris. Parinkus b koeficiento reikšmes, a koeficientų reikšmės gali būti apskaičiuotos mažiausiu kvadratų metodu. b koeficiento reikšmė didinama tol kol gaunamas mažiausias funkcijos (2.18) kvadratinis nuokrypis nuo ekperimentinių duomenų.

Biomasės koncentracijos kitimo grafikas aproksimuojamas iš dviejų dalių, nuo 0 iki 12 valandos trečiu laipsniu, o nuo 12 iki 13,870 valandos aproksimuojama pirmu laipsniu. Substrato grafikas aproksimuojamas iš trijų dalių – nuo 0 iki 8 valandų ir nuo 8 iki 12,865 trečiu laipsniu, o nuo 12,865 iki 13,870 valandos aproksimuojama pirmu laipsniu:



2.8 pav. biomasės koncentracijos kitimo grafikas laiko atžvilgiu aproksimuotas logistine funkcija.



2.9 pav. substrato koncentracijos kitimo grafikas laiko atžvilgiu aproksimuotas logistine funkcija.

2.2 Parametrų identifikavimas pirmu priartėjimu [11].

Norint surasti matematinio modelio parametrų reikšmes dažniausiai taikomi atsitiktinės paieškos metodo algoritmas. Svarbi identifikavimo algoritmų ypatybė – jų veiksmingumas labai priklauso nuo pradinių parametrų pasirinkimo. Todėl turint priminės informacijos apie galimas modelio parametrų reikšmes, tikslinga ją panaudoti. Norint gauti pakankamai gerus rezultatus ir sumažinti skaičiavimų laiką, pirminės parametrų reikšmės turi būti parinktos kuo artimesnės optimaliosioms reikšmėms, todėl šiame skyriuje bus identifikuojami parametrai pirmu priartėjimu. Iš anksto nustačius proceso kokybės kriterijų, reikia atlikti jautrumo analizę ir įvertinti, kurie modelio parametrai jam daro didžiausią ir kurie mažiausią poveikį.

Pirmu priartėjimu bus identifikuojami k_s , k_i , k_x , μ_{\max} ir Y_{xs} parametrai. Parametrai κ_1 , κ_2 , κ_3 ir κ_4 bus prilyginti nuliui. κ_1 , κ_2 parametrus prilyginus nuliui (1.17) išraišką galime išsireikšti taip:

$$\frac{s}{\mu+k_x} = \frac{k_s}{\mu_{\max}} + \frac{1}{\mu_{\max}} \cdot s + \frac{1}{k_i \cdot \mu_{\max}} \cdot s^2 \quad (2.19)$$

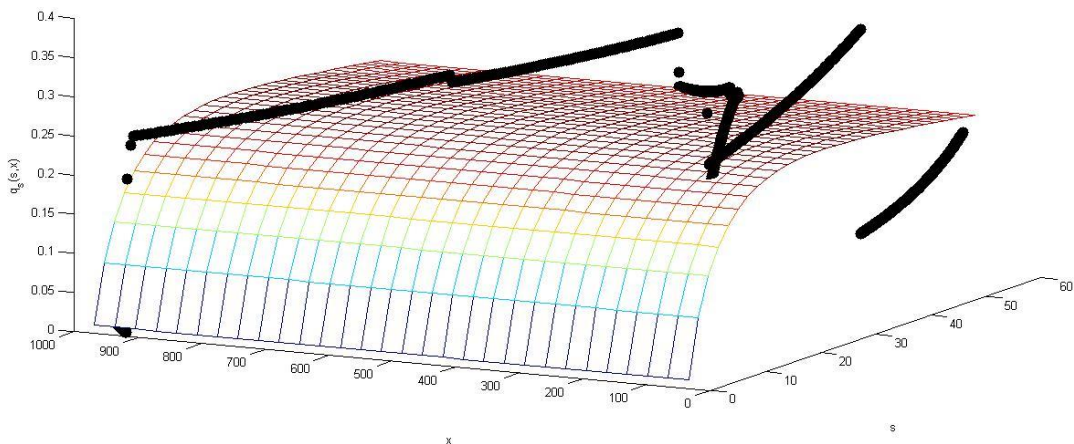
Turint aproksimuotus duomenis s_j , μ_j ir priskiriant reikšmę koeficientui k_x galime apskaičiuoti koeficientų k_s , k_i , μ_{\max} vertes mažiausių kvadratų metodu. Koeficiento k_x vertė bus nuosekliai didinama tol, kol modelis labiausiai atitiks aproksimuotų taškų vertes. Tai bus atliekama skaičiuojant minimalų vidutinį kvadratinį nuokrypį.

κ_3 , κ_4 parametrus prilyginus nuliui (1.20) išraišką galime išsireikšti taip:

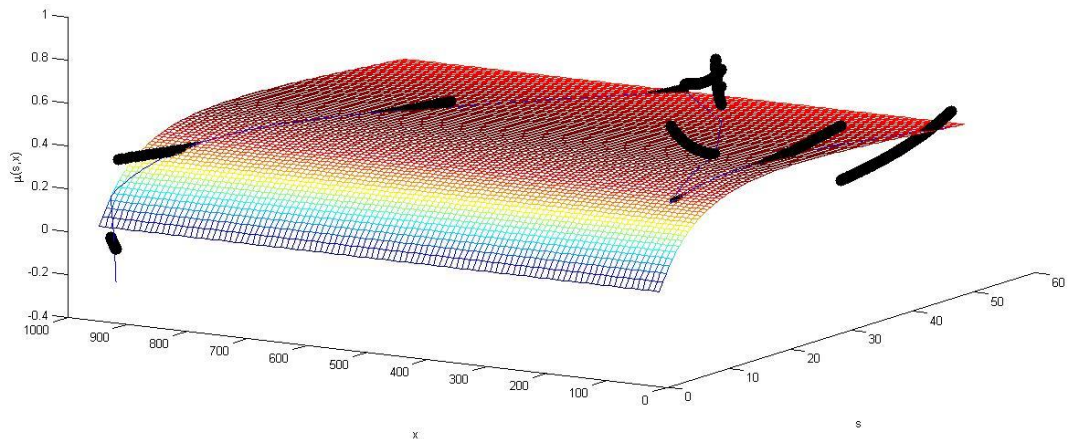
$$q(s, x) = Y_{xs} \cdot \mu_g(s, x) \quad (2.20)$$

Kadangi koeficientų k_s , k_i , μ_{\max} vertės jau yra žinomos, μ_g apskaičiuojamas iš (1.18) išraiškos. s_j ir q_j duomenys gaunami iš aproksimuotų eksperimentinių taškų. Y_{xs} randamas mažiausiu kvadratų metodu [12].

MATLAB aplinkoje nubrėžus biomasės ir substrato santykinų greičių priklausomybes nuo biomasės ir substrato koncentracijų galima palyginti kaip aproksimuoti duomenys atitinka matematinį modelį su pirminėmis parametru vertėmis. Modelyje κ_1 , κ_2 , κ_3 ir κ_4 parametru vertes prilyginus nuliui, biomasės koncentracijos vertės lieka nežinomos.



2.10 pav. substrato santykinio greičio priklausomybė nuo biomasės ir substrato koncentracijų



2.11 pav. biomasės santykinio greičio priklausomybė nuo biomasės ir substrato koncentracijų.

Turint visus skirtingais būdais aproksimuotus duomenis ir matematinį modelį, kurio parametrų vertės yra pirminės, galime palyginti aproksimuotų duomenų nuokrypį nuo pirminio modelio. Skaičiuojamas biomasės santykinio greičio nuokrypis nuo aproksimuotų duomenų:

$$\mu_{err} = \sum_{i=1}^n (\mu_{exp.i} - \mu_{sim.i})^2 \quad (2.21)$$

Skaičiuojamas substrato santykinio greičio nuokrypis nuo aproksimuotų duomenų:

$$q_{s,x,err} = \sum_{i=1}^n (q_{s,x,exp.i} - q_{s,x,sim.i})^2 \quad (2.22)$$

2.3 lentelė. Aproksimuotų duomenų nuokrypis nuo pirminio modelio.

Aproksimavimo būdas	$\mu(s,x)$ nuokrypis	$q(s,x)$ nuokrypis
Linijinė interpoliacija	14,6653	3,4804
Interpoliacija splineu	11,5085	3,4114
Ermito interpoliacija	11,4001	3,1047
Aproksimavimas logistine funkcija	17,8198	2,1666
Interpoliacija polinomu	14,5536	1,4862

Aproksimuojant biomasės koncentracijos duomenis Hermito interpoliacija ir perskaičiavus biomasės santykinį greitį $\mu(s,x)$ gaunamas mažiausias nuokrypis nuo matematinio modelio esant pirminėms parametrų reikšmėms, Tačiau aproksimavus substrato koncentracijos duomenis Hermito interpoliacija, nuokrypis gaunamas didesnis nei aproksimuojant logistine funkcija ir interpoliuojant polinomu.

2.3 Parametrų identifikavimas.

[11] [13] Modelio parametrus identifikuoti dažniausiai taikomi atsitiktinės paieškos metodai. Esant tam tikros tikslo funkcijos paviršiui gradientiniai ir tiesioginės paieškos metodai gali patekti į lokalinį minimumą. Šiuo atveju būtų tikslinga taikyti atsitiktinės paieškos metodą.

Netiesinio modelio parametrų identifikavimas apibrėžiamas taip: reikia rasti parametrų vektoriaus p reikšmes, kai būsenos kintamieji c aprašomi funkcine priklausomybe $f(t, p)$:

$$c = f(t, p) \quad (2.23)$$

ir turimi proceso eksperimentiniai duomenys $[t_e \ c_e]$. Darbe nagrinėjamu atveju $f(t, p)$ rezultatas yra biomasės arba substrato koncentracijų reikšmės.

Paprastai netiesinio optimizavimo metoduose algoritmas inicializuojamas pasirenkant pradinį parametrų vektorius p_0 . Toliau naudojant funkciją $f(t, p)$ apskaičiuojamos funkcijos reikšmės c laiko momentais t_e . Funkcija $c=f(t, p)$ gali būti algebrinė lygtis (polinomas, splainas, dirbtinis neuroninis tinklas) arba diferencialinė lygtis, tokių lygčių sistema arba algebrinių ir diferencialinių lygčių rinkinys (hibridinis modelis).

Parametrų paieškos procedūros tikslas - rasti tokias optimizuojamo parametrų vektoriaus p vertes, kurios minimizuotų modelio $c=f(t, p)$ nuokrypį nuo eksperimentinių duomenų $c_e(t_e)$. Taigi apskaičiuoto parametrų vektoriaus tinkamumas įvertinamas sudarant tikslo funkciją J , kurios išraiška gali būti apibrėžiama kaip vidutinis kvadratinis arba absoliutinis nuokrypis tarp modeliavimo ir eksperimentinių rezultatų.

$$J = \min \frac{\sum (c(t,p) - c(t))^2}{n} \quad (2.24)$$

Atsitiktinės paieškos algoritmai optimalaus parametrų vektoriaus paiešką pradeda p_0 aplinkoje. Kiekvienas naujai apskaičiuotas parametrų vektorius p , sumažinantis J reikšmę naudojamas tolesnėje paieškoje apibrėžiant naują paieškos sritį. Paieška nutraukiama pasiekus maksimalų paieškos algoritmo iteracijų skaičių, nustatytą J reikšmę arba kai santykinis J pagerėjimas nedidelis.

Vadinamasis brutalių jėgų (brutal force) atsitiktinės paieškos metodas atsitiktiniu būdu generuoja parametrų reikšmes ir iš jų atrenka geriausias. Atlikus statistiškai pakankamą kiekį iteracijų, randamas globalus tikslo funkcijos optimumas. Pagrindinis šio metodo pranašumas – jis paprastas ir užtikrina sėkmę. Skaičiavimams sugaištama daug laiko, tačiau tobulėjant informacinėms technologijoms skaičiavimo laikas vis mažėja. Atliekant daugelio gana sudėtingo modelio parametrų identifikavimą ir (arba) turint ribotą laiką skaičiavimams, šio metodo taikymas gali neduoti norimo rezultato.

Šios procedūros algoritmas primena mikroorganizmų elgseną ieškant substrato. Mikroorganizmai gali ieškoti substrato tik nedidelėje jų kūną supančioje srityje. Vos kurioje nors pusėje aptikęs substrato, mikroorganizmas juda ta kryptimi, sunaudoja surastą substratą ir toliau jo ieško naujai pasiektos srities aplinkoje. Todėl nagrinėjama optimizavimo procedūra vadinama Chemotaksis.

Šis atsitiktinės paieškos metodas parametrų p paiešką atlieka tam tikrame apie parametrų centrinių reikšmių tašką p_c apribotame intervale Δp . Pradžioje p_c priskiriamos pradinio parametrų vektoriaus p_0 reikšmės. RADIUS naują parametrų vektorius, pagerinantį $J(p)$ reikšmę, juo pakeičiamas centras p_c ir atliekama tolesnė atsitiktinė paieška.

Dar vienas būdas pagerinti atsitiktinės paieškos algoritmų spartą – naudoti evoliucijos procesą.

Evoliucinis programavimas buvo sugalvotas 1960 m. kaip dirbtinio intelekto alternatyva. Evoliucinis programavimas taikomas prognozėms, modeliavimui ir nežinomų procesų kontroliavimui [14]. Pagrindinė evoliucinio programavimo idėja paremta biologinių rūšių evoliucija. Evoliucija vyksta rūšies populiacijoje, t. y. individų aibėje. Šie individai generuoja naujus narius, o kiti (silpniausi) populiacijos nariai išnyksta. Generavimo proceso metu nauji nariai (palikuonys arba vaikai) sudaromi naudojant kitų narių (tėvų) informaciją (šiuo atveju – skaitines atitinkamų parametrų reikšmes). Vaikų požymiai arba savybės nustatomos naudojantis kryžminimo (angl. crossover) poveikiu ir (ar) mutacija (atsitiktinio pobūdžio keitimu). Šiame darbe bus naudojama mutacija be kryžminimo poveikio. Kaip ir gamtoje gyvuojančiose populiacijose, naujus narius kurti leidžiama geriausiems nariams (atsižvelgiant į pasirinktą kokybės kriterijų – optimizavimo procedūros tikslo funkciją). Jie paveldi geriausias tėvų savybes. Kita vertus, tik stipriausi nariai turi galimybę išgyventi, nes populiacijos dydis ribotas ir silpniausius iš populiacijos išstumia atsirandantys stipresni nariai.

Taikant šią idėją, iš esmės atsitiktinės paieškos algoritmas vykdomas taip pat, kaip jau aprašyta, tačiau generuojamas ir vienu metu išbandomas daugelis parametrų rinkinių derinių. Iš sudarytų parametrų rinkinių (populiacijos) pašalinami blogiausių tikslo funkcijos reikšmę užtikrinantys rinkiniai, o jų vietą užima nauji. Nauji rinkiniai generuojami geriausių rinkinių arba rinkinių derinių srityse. Nauji deriniai gali būti sudaromi kombinuojant (kryžminant) jau turimų geriausių parametrų vektorių reikšmes (angl. - crossover act).

Būtina pabrėžti, kad netgi naudojant paprasčiausius atsitiktinės paieškos algoritmus, biotechnologinių procesų modelių parametrų sprendinius galima rasti pakankamu tikslumu. Sudėtingesni algoritmai gali tik paspartinti paieškos procedūrą.

Optimizavimo uždavinys bendriausioje formoje formuluojamas taip [15]:

$$J = f(p) \rightarrow \min(\max), \quad p = [x_1, \dots, x_n]^T \quad (2.25)$$

Algoritmo realizacija pradedama nuo kintamųjų vektorių pradinės sekos parinkimo:

$$P_1, P_2, \dots, P_p.$$

Ši seka dar vadinama *vektoriai-tėvai* (angl. *parents*). Konkrečiu atveju vektorius-tėvas yra vektorius sudarytas iš darbe naudojamo matematinio modelio koeficientų μ_{\max} , k_s , k_i , k_x , Y_{xs} , κ_1 , κ_2 , κ_3 , κ_4 reikšmių.

Toliau atliekama šios pradinės sekos “mutacija”, kurios rezultate gaunamos naujos kintamųjų kombinacijos - *vektoriai-palikuonys* (angl. *offsprings*).

Jei paimsime i -ąjį vektorių p (palikuonys), jo “mutacija” atliekama taip:

$$p_i(\text{palikuonys}) = \mathbf{x}_{i_n} + \Delta \mathbf{x}_i \cdot \text{rand}_i,$$

t.y.,

$$p_i(\text{palikuonys}) = \begin{bmatrix} x_{i1} \dots x_{in} + \Delta x_1 \cdot \text{rand}_1 \\ x_{i1} \dots x_{in} + \Delta x_2 \cdot \text{rand}_2 \\ \vdots \\ x_{i1} \dots x_{in} + \Delta x_i \cdot \text{rand}_i \end{bmatrix}, \quad i = 1, \dots, p, \quad (2.26)$$

čia $RAND$ - atsitiktinis Gauso kintamasis su nuline matematine viltimi $M\{RAND\} = 0$ ir vienetine dispersija $D\{RAND\} = 1$.

Δx_n - parametrai, nustatantys “mutacijų” laipsnį.

Skaičiavimų algoritmuose tai yra derinimo parametrai, turintys įtakos algoritmų konvergavimo greičiui.

Nauja *vektorių-tėvų* seka sudaroma iš ankstesnės *vektorių-tėvų* ir *vektorių-mutantų* sekos, atrenkant vektorius, kurie atitinka kokybės kriterijų – didžiausias arba mažiausias tikslo funkcijos vertes priklausomai nuo tikslo funkcijos ir norimo gauti rezultato.

Algoritmas atlieka šią atranką, surikiuodamas visus vektorius-tevus ir vektorius-palikuonis nuo geriausio iki blogiausio pagal tikslo funkcijos vertę. Atrenkamas nustatytas geriausių vektorių skaičius sekančių “mutantų” generavimui.

Ši procedūra baigiama, kai pasiekiamas ekstremali tikslo funkcijos vertė (kai nebegaunama tikslo funkcijos pagerėjimo) arba kai pasiekiamas užduotas iteracijų skaičius.

Pagrindiniai metodo privalumai:

- universalumas,
- paprastas realizavimas,
- galimybė rasti globalinį ekstremumą.

Pagrindinis trūkumas:

- didelis iteracijų skaičius, reikalingas ekstremumo nustatymui.

Apibendrinant, algoritmo realizavimo žingsniai tokie:

Parinkti pradinę p_0 vektorių-tėvų populiaciją p_1, p_2, \dots, p_p . Parametrų μ_{\max} , k_s , k_i , k_x , Y_{xs} vertės buvo apskaičiuotos identifikuojant parametrus pirmu priartėjimu. Pradinės parametrų $\kappa_1, \kappa_2, \kappa_3$ ir κ_4 vertės, kurios prieš tai buvo prilygintos nuliui, parenkamos visiems vienodos ir lygios pirmu priartėjimu apskaičiuoto mažiausio koeficiento vertei ir padaugintai iš 0,1. Pradinės p vektorių-tėvai p_1, p_2, \dots, p_p bus visi vienodi. Pvz.:

$$p_i(\text{tėvai}) = \begin{bmatrix} p_1 = [1,0043 & 6,7567 & 91,1183 & 0,2388 & 0,02 & 0,02 & 0,02 & 0,3782 & 0,02] \\ p_2 = [1,0043 & 6,7567 & 91,1183 & 0,2388 & 0,02 & 0,02 & 0,02 & 0,3782 & 0,02] \\ p_p = [1,0043 & 6,7567 & 91,1183 & 0,2388 & 0,02 & 0,02 & 0,02 & 0,3782 & 0,02] \end{bmatrix} \quad (2.27)$$

1. Iš kiekvieno „tėvų“ vektoriaus p_i , $i=1, \dots, p$ sudaryti atitinkamus palikuonių vektorius p_i , $i=p+1, \dots, 2p$, pridedant prie kiekvieno vektoriaus elementų atsitiktinius Gauso kintamuosius su nuliniu vidurkiu ($M=0$) ir normuota dispersija ($D=1$), padaugintus iš parametrų Δx_i , nustatančių „mutacijų“ laipsnį.

$$p_i(\text{palikuonys}) = \begin{bmatrix} p_{1+p} = [x_{1\dots} x_n + \Delta x \cdot \text{rand}_1] \\ p_{2+p} = [x_{2\dots} x_n + \Delta x \cdot \text{rand}_2] \\ p_{p+p} = [x_{p\dots} x_n + \Delta x \cdot \text{rand}_p] \end{bmatrix} \quad (2.28)$$

2. Apskaičiuoti tikslo funkcijos vertes vektoriams-tėvams $f(p_i)$, $i=1, \dots, p$ ir vektoriams-palikuoniams $f(p_i)$, $i=p+1, \dots, 2p$. Sudaroma tikslo funkcija, kuri įvertina modeliavimo rezultatų minimalų kvadratinį suminį nuokrypį (koncentracijose x ir s) nuo aproksimuotų eksperimentinių proceso duomenų:

$$J = \min \sum_{j=1}^N \left(\left(\frac{(x_j - x_{mj})}{x_{vid}} \right)^2 + \left(\frac{(s_j - s_{mj})}{s_{vid}} \right)^2 \right) \quad (2.29)$$

x_j, s_j – aproksimuotos eksperimentinės reikšmės;

x_{mj}, s_{mj} – reikšmės po mutacijos;

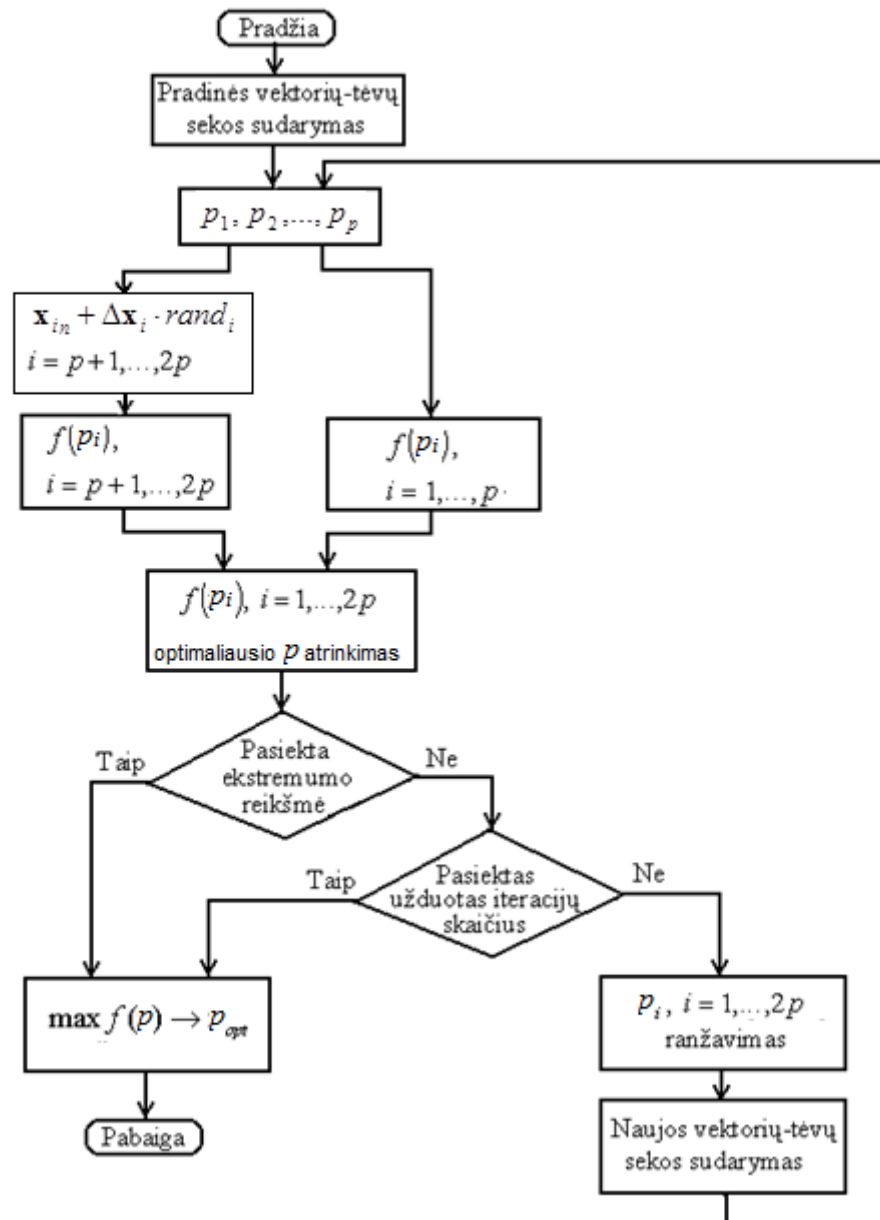
x_{vid}, s_{vid} – aproksimuotų eksperimentinių reikšmių vidurkis;

5. Atliekamas p vektorių surikiavimas pagal apskaičiuotą mažiausią tikslo funkcijos vertę. Išrenkamas nustatytas skaičius vektorių, kurių tikslo funkcijos vertės yra mažiausios ir su tais vektoriais yra atliekama sekanti iteracija.
6. Skaičiavimas baigiamas, kai mutacijos nebeduoda tikslo funkcijos pagerėjimo, arba pasiekiamas nustatytas iteracijų skaičius.

Kitu atveju iteracinis skaičiavimas tęsiamas, pereinant prie 2-o žingsnio.

7. Pasiekus nustatytą iteracijų skaičių mažinamas mutacijos laipsnis iki $\frac{\Delta x_n}{7}$. Atlikus skaičiavimus nustatytą iteracijų skaičių su pakeistu mutacijos laipsniu, mutacijos laipsnis dar kartą mažinamas iki $\frac{\Delta x_n}{50}$. Atlikus visas iteracijas su sumažintu mutacijos laipsniu gaunamos optimalios p_i vektoriaus koeficientų vertės.

Algoritmo struktūra:



2.12 pav. chemotaxis ir evoliucinio metodų derinio algoritmas

2.4 Pamaitinimo profilio optimizavimas.

Daroma prielaida, kad yra tam tikra santykinio substrato sunaudojimo greičio reikšmė, kurią viršijus substratas yra naudojamas ne biomasės gamybai, o šalutinio produkto gamybai, kuris lėtina biomasės gamybą. Todėl yra svarbu nustatyti tinkamą pamaitinimo greitį maksimaliai biomasės išeigai gauti.

Vadovaujantis skirtingų formų pamaitinimo profilių tyrimo rezultatais [16] buvo pasirinktas eksponentinio tipo pamaitinimo profilis, kuris reikalauja nedaug parametrų ir gerai aproksimuoja optimalų pamaitinimą:

$$u(t) = \begin{cases} 0, & \text{if } t < t_{feed-start}; \\ f(t) = \alpha \exp(\beta(t - t_{feed-start})), & \text{if } t_{feed-start} \leq t \leq t_{feed-end}; \\ 0, & \text{if } t > t_{feed-end}; \end{cases} \quad (2.30)$$

$t_{feed-start}$, $t_{feed-end}$, α ir β parametrai reikalingi optimizavimui atlikti.

1. Parinkti pradinę p vektorių-tėvų populiaciją p_1, p_2, \dots, p_p . Pamaitinimo pradžios laiko $t_{feed-start}$ pirminė vertė parenkama iš eksperimentinių pamaitinimo duomenų - $t_{feed-start} = 8$, α pirminis santykinis pamaitinimo greitis laiko momentu $t_{feed-start}$ parenkamas iš pamaitinimo eksperimentinių duomenų ir lygus $\alpha = 0,2$, β parametras, įvertinantis eksponentės išlenktumą parenkamas taip pat iš eksperimentinių duomenų. Pradiniai vektoriai-tėvai visi vienodi, jų skaičius priklauso nuo nustatyto mutacijų skaičiaus.

$$p_i(\text{tėvai}) = \begin{bmatrix} p_1 = [8 & 0,2 & 0,2] \\ p_2 = [8 & 0,2 & 0,2] \\ p_p = [8 & 0,2 & 0,2] \end{bmatrix} \quad (2.31)$$

2. Iš kiekvieno „tėvų“ vektoriaus p_i , $i=1, \dots, p$ sudaryti atitinkamus palikuonių vektorius p_i , $i=p+1, \dots, 2p$, pridendant prie kiekvieno vektoriaus elementų atsitiktinius Gauso kintamuosius su nuliniu vidurkiu ($M=0$) ir normuota dispersija ($D=1$), padaugintus iš parametrų Δx_n , nustatančių „mutacijų“ laipsnį.

$$p_i(\text{palikuonys}) = \begin{bmatrix} p_{p+1} = [8 & 0,2 & 0,2] + \Delta x \cdot rand_1 \\ p_{p+2} = [8 & 0,2 & 0,2] + \Delta x \cdot rand_2 \\ p_{2p} = [8 & 0,2 & 0,2] + \Delta x \cdot rand_p \end{bmatrix} \quad (2.32)$$

3. Atlikus 1 ir 2 veiksmus gauname eksponentinę kreivę, kuri atitinka spėjamo santykinio pamaitinimo greičio kitimą. Su kiekvienu p vektoriumi pagal medžiagų masės balanso lygtis ir apskaičiuotą spėjamą santykinio pamaitinimo greičio kitimą apskaičiuojama ODE113 aprašyti kaip skaičiuoja. Susistatom reikšmes ir gauname rezultatą.

4. Atlikus 1 ir 2 veiksmus gauname eksponentinę kreivę, kuri atitinka spėjamo santykinio pamaitinimo greičio kitimą. Naudojantis MATLAB funkcija diferencialinėms lygtims spręsti ODE113 su kiekvienu p vektoriumi pagal medžiagų masės balanso lygtis ir apskaičiuotą spėjamą santykinio pamaitinimo greičio kitimą, apskaičiuojamos tikslo funkcijos vertės vektoriams-tėvams $f(p_i)$, $i=1, \dots, p$ ir vektoriams-palikuoniams $f(p_i)$, $i=p+1, \dots, 2p$. Funkcija ODE113 taiko kintamos eilės Adamso, Bašforto ir Moultono metodą diferencialinėms lygtims spręsti. Tikslo funkcija, kuri įvertina maksimalią biomasės išėigą:

$$J = \max V(t_e) x(t_e) \quad (2.33)$$

t_e – biotechnologinio proceso pabaigos laikas.

5. Atliekamas p vektorių surikiavimas pagal apskaičiuotą didžiausią tikslo funkcijos vertę. Išrenkamas nustatytas skaičius vektorių, kurių tikslo funkcijos vertės yra didžiausios ir su tais vektoriais yra atliekama sekanti iteracija.
6. Skaičiavimas baigiamas, kai mutacijos nebeduoda tikslo funkcijos pagerėjimo, arba pasiekiamas nustatytas iteracijų skaičius.
- Kitu atveju iteracinis skaičiavimas tęsiamas, pereinant prie 2-o žingsnio.
7. Mutacijos laipsnis Δx šiuo atveju pastovus.

3 Biotechnologinių procesų matematiniai modeliavimo rezultatai MATLAB aplinkoje.

3.1 Parametrų vertės pirmu priartėjimu:

3.1 lentelė. Parametrų vertės pirmu priartėjimu.

Aproksimavimo metodas	Parametrų vertės								
	μ_{\max}	k_s	k_i	k_x	κ_1	κ_2	κ_3	Y_{xs}	κ_4
Linijinė interpoliacija	1,0043	6,7567	91,1183	0,2388	0	0	0	0,3782	0
Interpoliacija spline	1,4720	13,2658	37,5428	0,2243	0	0	0	0,3888	0
Ermito interpoliacija	1,3280	12,2644	44,2110	0,2130	0	0	0	0,3971	0
Aproksimavimas logistine funkcija	0,8197	17,0665	9484,0	0,1188	0	0	0	0,4641	0
Interpoliacija polinomu	0,723	5,5017	54620	0,2029	0	0	0	0,3917	0

3.2 Parametrų optimizavimas chemotaksis ir evoliucinio metodų deriniu (mutacijų skaičius 10)

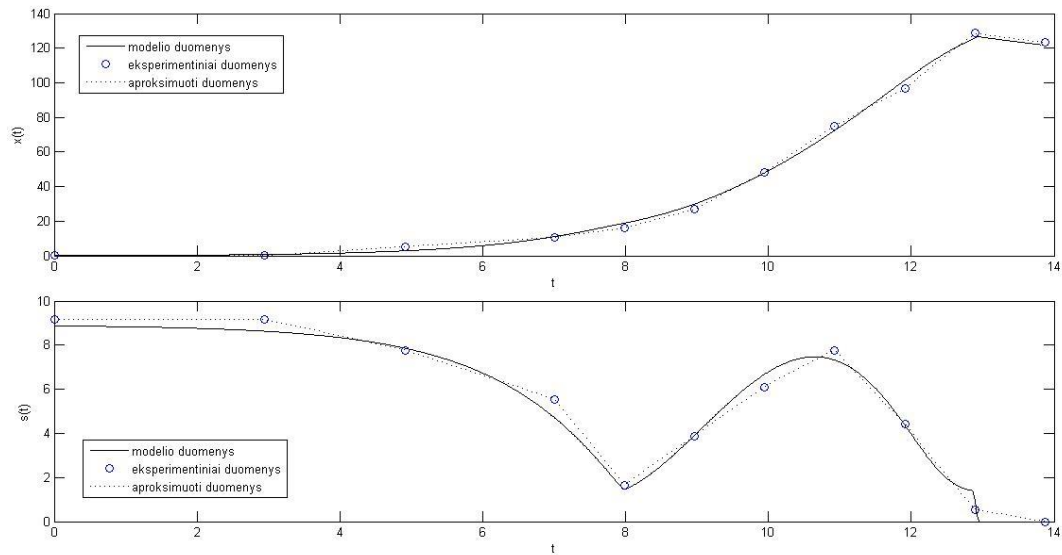
3.2 lentelė. Parametrų optimizavimas chemotaksis ir evoliucinio metodų deriniu (mutacijų skaičius 10).

Aproksimavimo metodas	Iteracijų skaičius	nuokrypis	Parametrų vertės								
			μ_{\max}	k_s	k_i	k_x	κ_1	κ_2	κ_3	Y_{XS}	κ_4
Linijinė	30	2412,8	2,3204	5,7524	122,1780	0,1526	0,0250	0,0127	0,0245	0,4326	0,0227
	150	211,2	1,6438	6,1361	94,7558	0,0986	0,0050	0,0027	0,0150	0,1056	0,0037
	450	22,4611	1,0168	2,7047	105,0198	0,0146	0,0041	0,0007	0,0434	0,2721	0,0022
	1500	8,0391	0,8294	0,8760	99,3370	0,0188	0,0062	0,0002	0,0527	0,2757	0,0024
Splinas	30	2347,4	2,5360	6,0471	32,3182	0,1550	0,0139	0,0168	0,0207	0,3253	0,0214
	150	192,2	1,6609	2,8124	10,4307	0,0616	0,0122	0,0011	0,0162	0,1926	0,0063
	450	37,8748	1,7839	5,3932	11,0145	0,0281	0,0082	0,0005	0,0108	0,2302	0,0044
	1500	14,9525	0,8976	1,014	46,666	0,0235	0,0053	0,0004	0,0108	0,3468	0,0023
Ermito	30	2663,9	2,2969	6,4879	48,9250	0,1418	0,0184	0,0184	0,0175	0,4860	0,0414
	150	269,0280	1,8630	7,7018	48,7901	0,0593	0,0109	0,0018	0,0223	0,1012	0,0063
	450	29,5671	1,2724	3,9109	31,6843	0,0333	0,0034	0,0009	0,0221	0,3115	0,0019
	1500	11,0738	1,2747	1,9499	12,9209	0,0298	0,0040	0,0008	0,0426	0,3271	0,0011
Logistinė	30	3174,6280	1,8501	12,3712	15038,3476	0,0864	0,0135	0,0106	0,0173	0,3357	0,0116
	150	183,0819	1,6129	7,0017	15624,7884	0,0617	0,0069	0,0019	0,0098	0,1591	0,0045
	450	19,1209	0,9200	2,3261	2462,9880	0,0186	0,0032	0,0006	0,0231	0,3419	0,0018

„Duomenų aproksimavimo metodų taikymas biotechnologinių procesų modeliavimui“

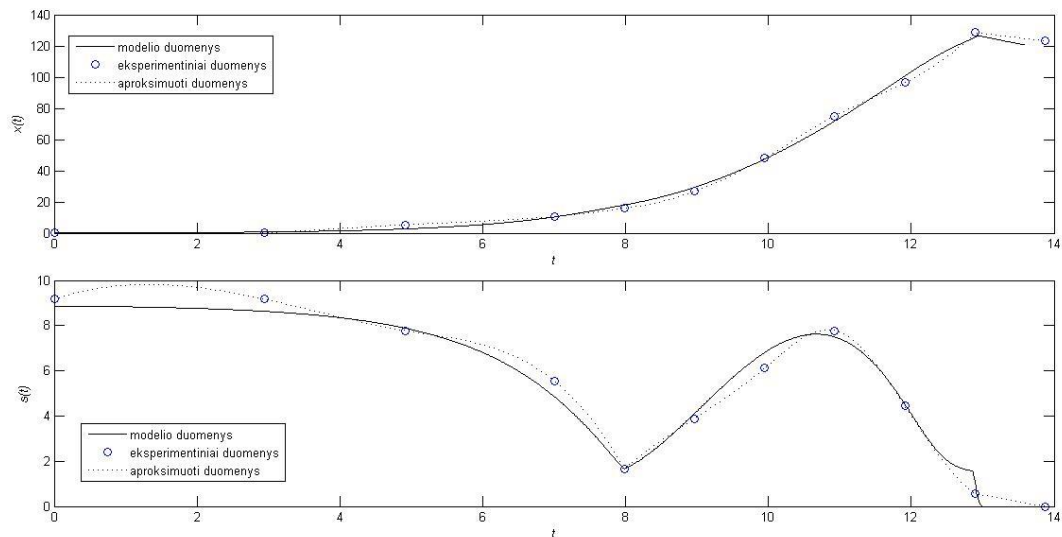
	1500	6,4279	0,0007	0,0007	6,5060	0,00001	0,000003	0,0000006	0,000008	0,0003	0,000001
Polinomu	30	2348,1004	1,2074	4,1540	59167,8730	0,0854	0,0137	0,0106	0,0213	0,3575	0,0231
	150	292,3986	0,9918	1,5693	42011,4647	0,1024	0,0085	0,0018	0,0229	0,1243	0,0067
	450	17,9949	0,9205	1,8571	47110,7644	0,0464	0,0047	0,0005	0,0109	0,3102	0,0024
	1500	8,0543	0,00007	0,00006	6,4286	0,000002	0,0000005	0,00000004	0,000009	0,00002	0,0000001

Atliekant skaičiavimus buvo pastebėta, jog esant 30 - 450 iteracijų skaičiui, apskaičiuotas nuokrypis vieną kartą gaunamas mažesnis aproksimuojant vienokiu būdu, kitą kartą – kitokiu. Tai reiškia, jog iteracijų ir (ar) mutacijų skaičius yra nepakankamas, todėl buvo nuspręsta padidinti iteracijų skaičių iki 1500.



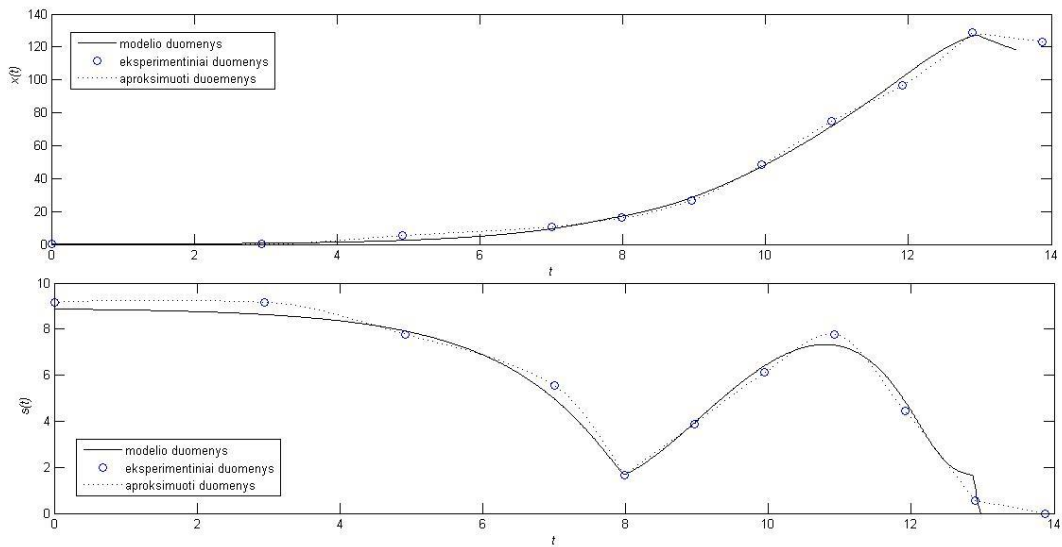
3.1 pav. optimizuotas modelis naudojant linijine interpoliacija.

Aproksimuojant linijine interpoliacija modelio nuokrypis nuo aproksimuotų verčių gaunamas pakankamai mažas 8,0391.



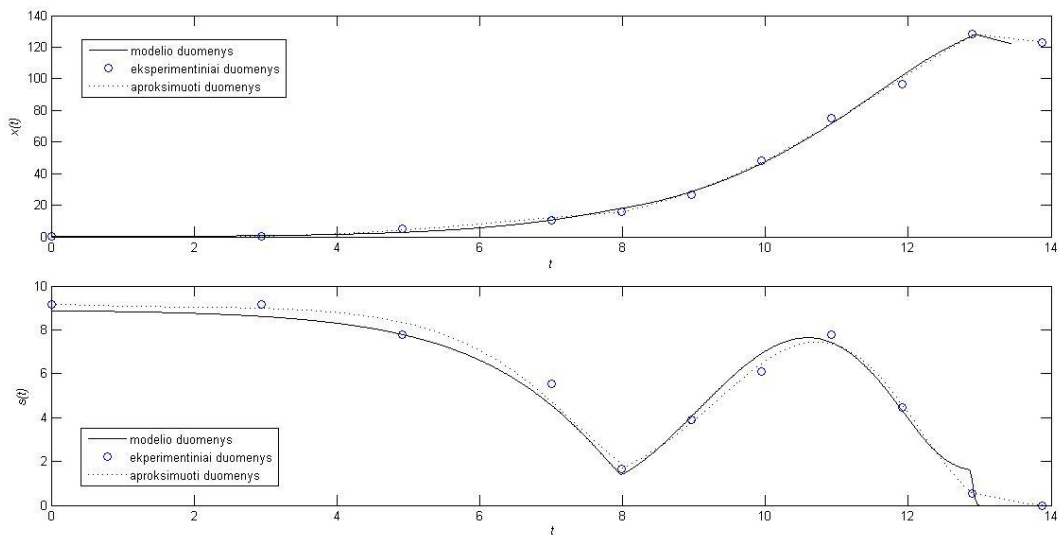
3.2 pav. optimizuotas modelis interpoliuojant splainu.

Aproksimuojant splainu modelio nuokrypis nuo aproksimuotų verčių gaunamas didžiausias 14,9525. Nuokrypis išaugo dėl akivaizdžiai matomo substrato padidėjimo ties 1 val. 30 min., kas yra neįmanoma, nes substratas turi mažėti iki pamaitinimo pradžios.



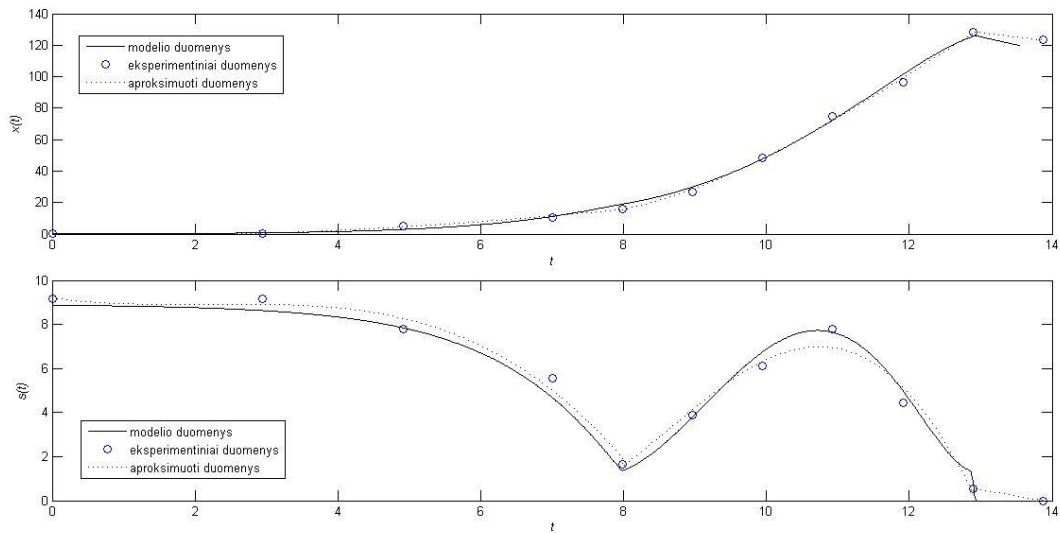
3.3 pav. optimizuotas modelis naudojant Hermito interpoliaciją.

Aproksimuojant kubine Hermito interpoliacija modelio nuokrypis nuo aproksimuotų verčių gaunamas 11,0738. Nuokrypis mažesnis nei aproksimuojant splainu, bet didesnis nei tiesiškai interpoliuojant.



3.4 pav. optimizuotas modelis naudojant aproksimavimą logistine funkcija.

Aproksimuojant dalimis 3 laipsnio logistine funkcija modelio nuokrypis nuo aproksimuotų verčių gaunamas mažiausias 6,4279. Galima pastebėti, kad šis aproksimavimo būdas skiriasi nuo kitų tuo, kad aproksimuojanti kreivė nekerta visų eksperimentinių taškų. Daroma prielaida, kad eksperimentiniai taškai pamatuoti esant tam tikrai paklaidai, kurią logistinė funkcija eliminuoja.



3.5 pav. optimizuotas modelis naudojant aproksimavimą polinomu.

Aproksimuojant dalimis 3 laipsnio polinomu modelio nuokrypis nuo aproksimuotų verčių gaunamas 0,0152 didesnis nei tiesiškai interpoliuojant. Kaip ir aproksimuojant logistine funkcija, aproksimuojant polinomu aproksimuojanti kreivė nekerta visų eksperimentinių taškų. Didžiausias skirtumas nuo aproksimavimo logistine funkcija, jog polinomu aproksimuojanti kreivė yra toliau nuo substrato pamaitinimo koncentracijos maksimumo, nei aproksimuojant logistine funkcija.

Išvados:

Aproksimavus eksperimentinius taškus skirtingais būdais, identifikavus modelio parametrus ir palyginus aproksimuotų taškų vertes su apskaičiuotomis modelio taškų vertėmis mažiausias nuokrypis (6,4279) esant maksimaliam iteracijų skaičiui nustatytas aproksimuojant logistine funkcija.

3.3 Pamaitinimo profilio optimizavimas:

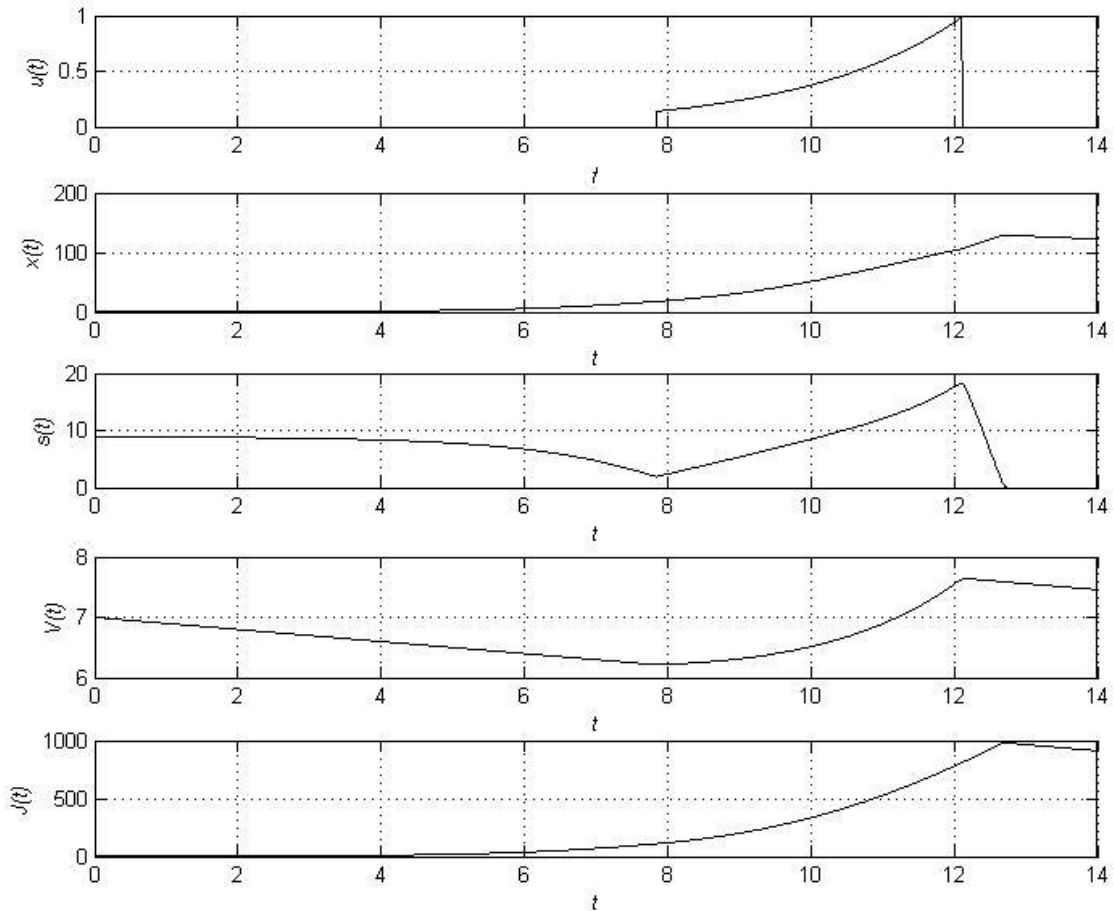
3.3 lentelė. Pamaitinimo profilio optimizavimo rezultatai.

Aproksimavimo metodas	Iteracijų skaičius (iteracijų skaičius identifikuojant parametrus)	Maksimali biomasės išeiga	Parametrų vertės		
			$t_{\text{feed_start}}$	α	β
Linijinė	300 (30)	299.2706	1.6668	0.3242	0.4054
	600 (150)	911.8007	2.6140	0.1107	0.3098
	900(450)	1011.3	5.5236	0.1013	0.4337
	1800(1500)	980.77	7.8403	0.1395	0.4578
Splainas	300 (30)	281.7425	2.7175	0.1220	0.4524
	600 (150)	714.1502	7.4831	0.1244	0.3732
	900(450)	910.5073	6.7165	0.0765	0.4679
	1800(1500)	969.4157	6.8533	0.0794	0.4298
Ermito	300 (30)	224.2443	2.6979	0.2780	0.2145
	600 (150)	790.5044	1.3480	0.0265	0.4959
	900(450)	992.2333	5.4124	0.0641	0.4235
	1800(1500)	981.6508	7.4900	0.1030	0.4372
logistinė	300 (30)	384.7116	0.2574	0.3394	0.3675
	600 (150)	939.5366	1.6069	0.2719	0.2589
	900(450)	1045.4365	5.6178	0.2726	0.4027
	1800(1500)	1006.5480	6.0515	0.2760	0.2869
Polinomu	300 (30)	308.4500	0.4454	0.1788	0.2494
	600 (150)	735.3159	4.1886	0.2577	0.3754
	900(450)	1018.2581	6.9341	0.4266	0.2767
	1800(1500)	1020.7285	5.8208	0.2081	0.3202

Matome, jog esant didesniai iteracijų skaičiui aproksimuojant tiesine interpoliacija, splainu ir logistine funkcija tikslinio produkto išeiga sumažėjo. Taip nutiko todėl, kad 3.2 poskyryje modelio

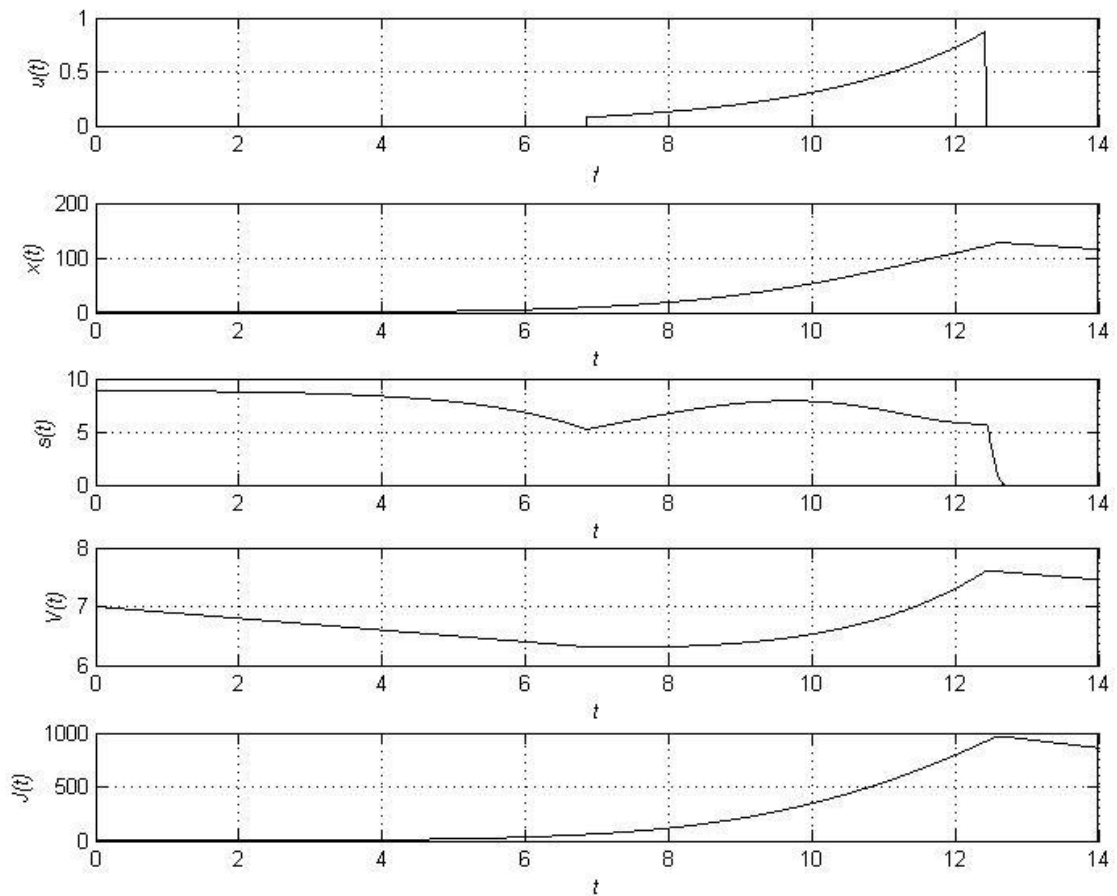
„Duomenų aproksimavimo metodų taikymas biotechnologinių procesų modeliavimui“

parametrai buvo atrinkti esant mažesniai iteracijų skaičiui ir su didesniu nuokrypiu nuo aproksimuotų duomenų. Žemiau pateikiami grafikai, kai modelio parametrai identifikuojami 1500 iteracijų skaičiumi, o pamaitinimo profilio parametrai optimizuojami 1800 iteracijų skaičiumi:



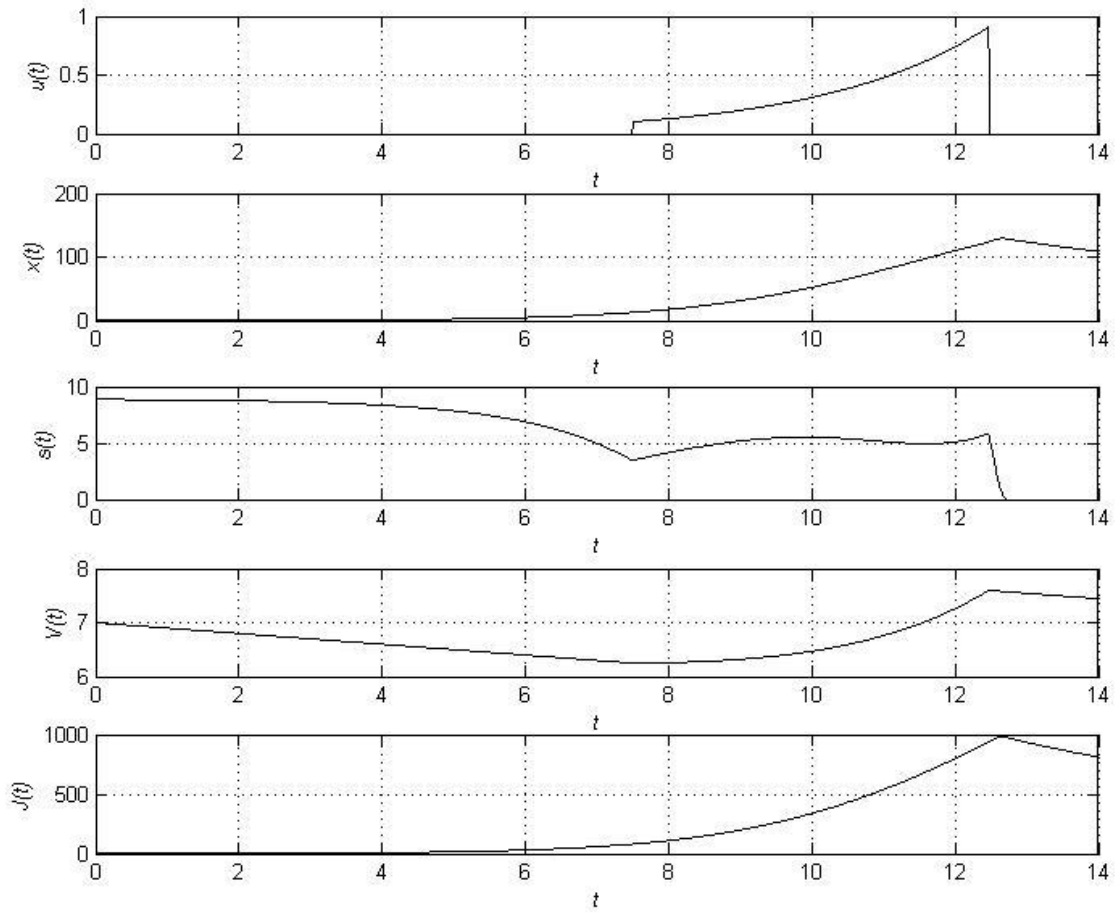
3.6 pav. optimizuotas pamaitinimo profilis tiesiškai interpoliuojant

Optimizuojant pamaitinimo profilį tiesiškai interpoliuojant tikslinio produkto išeiga gaunama 980,77 g. Biomės išeigos maksimumas pasiekiamas 12,7 h.



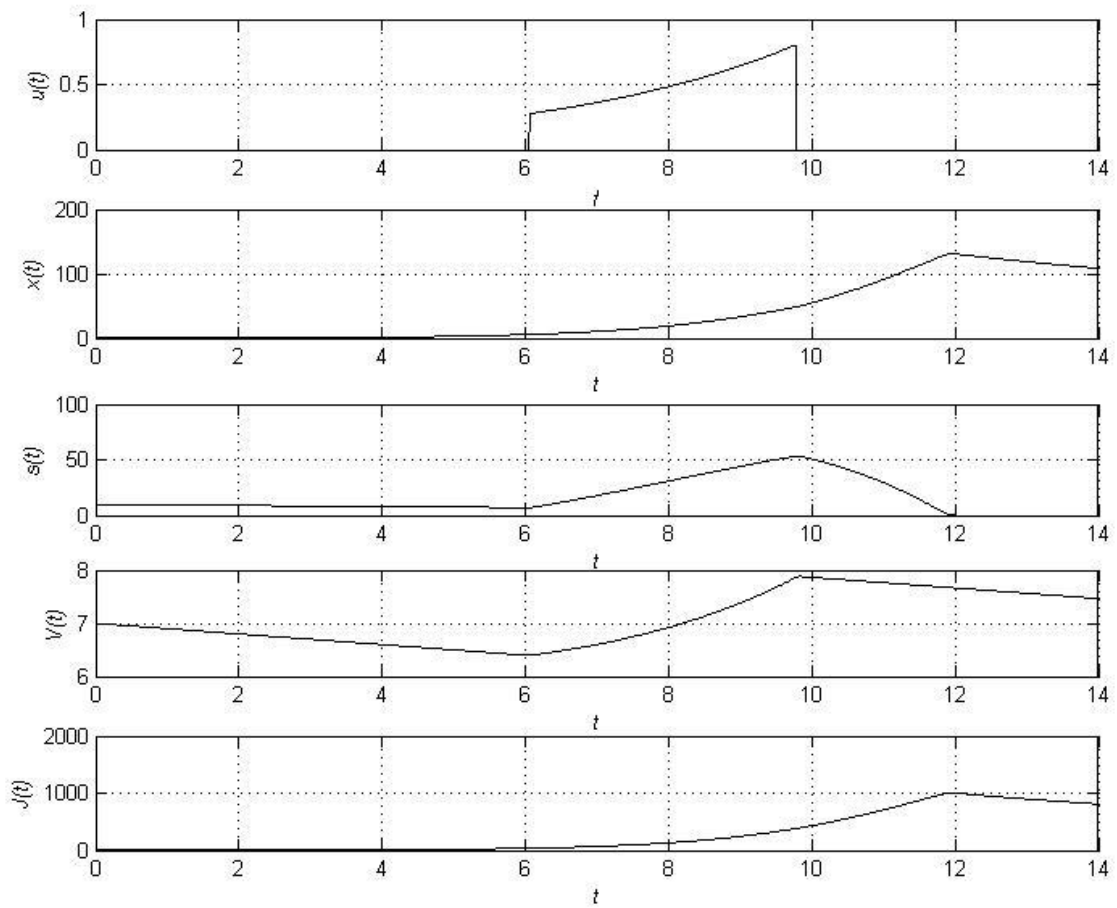
3.7 pav. optimizuotas pamaitinimo profilis interpoliuojant splainu

Optimizuojant pamaitinimo profilį splainu tikslinio produkto išeiga gaunama mažiausia 969,4157 g. Biomės išeigos maksimumas pasiekiamas 12,6 h.



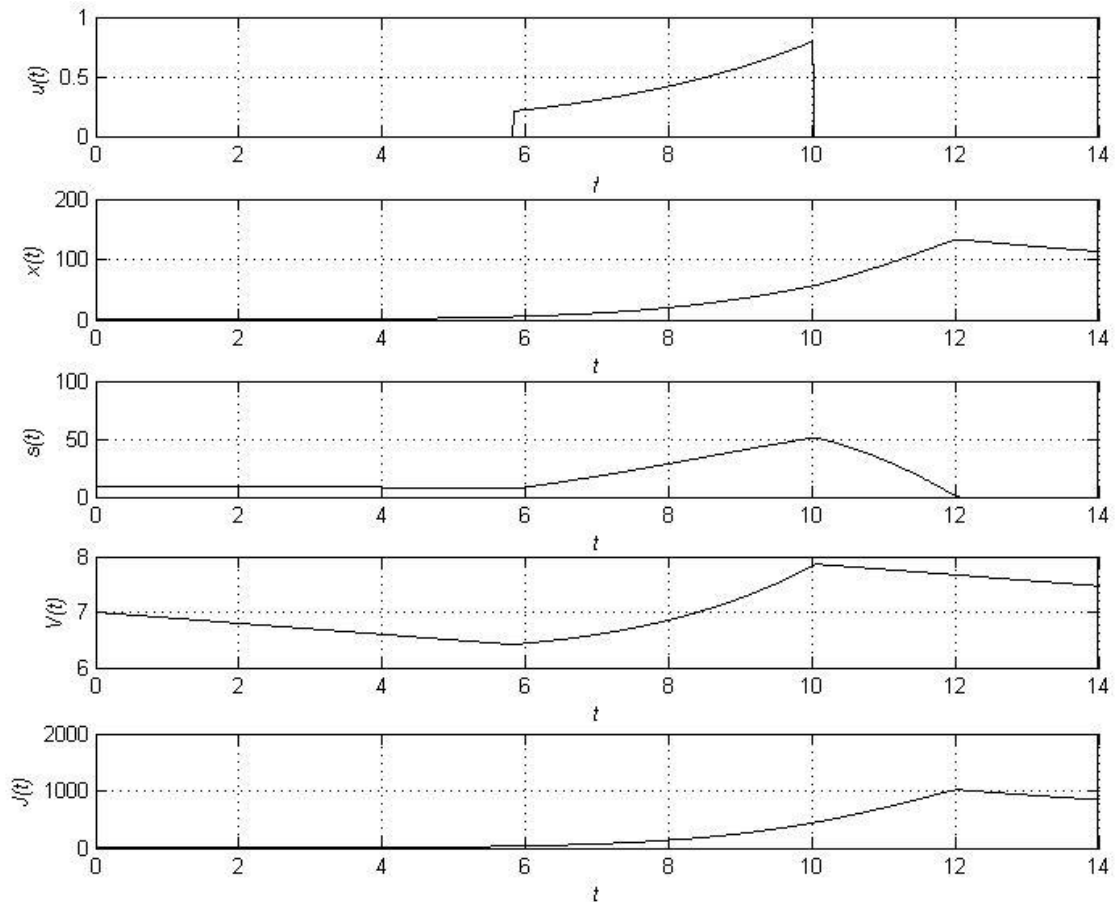
3.8 pav. optimizuotas pamaitinimo profilis naudojant Hermo interpoliaciją

Optimizuojant pamaitinimo profilį aproksimuojant Hermo interpoliacija tikslinio produkto išeiga gaunama 981.6508 g. Biomasės išeigos maksimumas pasiekiamas 12,7 h.



3.9 pav. optimizuotas pamaitinimo profilis aproksimuojant logistine funkcija

Optimizuojant pamaitinimo profilį aproksimuojant logistine funkcija tikslinio produkto išeiga gaunama 1006.5480 g. Biomasės išeigos maksimumas pasiekiamas greičiausiai – ties 12 h.



3.10 pav. optimizuotas pamaitinimo profilis interpoliuojant polinomu

Optimizuojant pamaitinimo profilį aproksimuojant polinomu tikslio produkto išeiga gaunama didžiausia 1020.7285 g. Biomasės išeigos maksimumas pasiekiamas 12,1 h.

Išvada:

Naudojant matematinį modelį, kuris buvo optimizuotas naudojant aproksimavimą polinomu, buvo gautas didžiausias tikslo funkcijos (maksimalios išeigos) rezultatas. Biomasės išeigos maksimali vertė buvo pasiekta per mažiausią laiko tarpą aproksimuojant logistine funkcija.

4 Išvados.

1. Darbe apžvelgta literatūra apie biotechnologinius procesus, bioreaktorių tipus ir jų skirtumus, biotechnologinių procesų matematinius modelius, jų parametrų identifikavimo metodus ir išanalizuoti eksperimento duomenų aproksimavimo metodai.

2. Atliekant tyrimą buvo suderinta programa MATLAB aplinkoje. Programa aproksimuoja eksperimento duomenis penkiais būdais: tiesine interpoliacija, polinomu, splineu, Ermito ir logistine funkcija. Identifikuoja biotechnologinio proceso modelio parametrus dviem etapais. Pirmuoju etapu mažiausiai įtakojantys modelį parametrai prilyginami nuliui, o likusių parametrų vertės randamos mažiausiu kvadratų metodu. Antruoju etapu optimizuojami modelio parametrai chemotaksis ir evoliucinio metodų deriniu. Pamaitinimo profilis buvo optimizuotas taip pat chemotaksis ir evoliucinio metodų deriniu.

3. Aproksimuojant eksperimentinius duomenis dalimis logistine funkcija gaunamas mažiausias nuokrypis (6.4279) nuo aproksimuotų duomenų iki modeliuojamų verčių. Modelio parametrai identifikuoti atlikus 1500 iteracijų po 10 mutacijų. Skaičiuojant modeliuojamų verčių nuokrypį nuo aproksimuotų verčių, abi biomasės ir substrato koncentracijų vertės buvo aproksimuojamos tuo pačiu metodu.

Optimizuojant pamaitinimo profilį, maksimali tikslinio produkto išeiga buvo didžiausia (1020.7285 g) aproksimuojant dalimis 3 laipsnio polinomu, nors ir modelio nuokrypis nuo aproksimuotų koncentracijos verčių buvo didesnis nei aproksimuojant logistine funkcija. Biomasės išeigos maksimumas pasiekiamas greičiau aproksimuojant logistine funkcija (12 h), nei aproksimuojant polinomu (12,1 h).

Tačiau maksimali biomasės išeiga nenusako modelio tikslumo. Modelio tikslumą įvertina aproksimuotų duomenų nuokrypis nuo modeliuojamų duomenų, šiuo atveju tiksliausias aproksimavimo būdas yra aproksimavimas dalimis 3 laipsnio logistine funkcija.

5 Literatūra.

1. Vytautas Galvanauskas, Donatas Levišauskas. Biotechnologinių procesų modeliavimas, optimizavimas ir valdymas. Vilniaus pedagoginio universiteto leidykla, 2008 m.;
2. J. Nielsen, J. Villadsen, G Liden. Bioreaction Engineering Principles. Kluwer Academic/ Plenum Publishers, New York, 2003 m.;
3. V. V. Biriukov, V. M. Kantere. Optimization of batch processes of microbiological synthesis. Nauka, Moscow, 1985 m.;
4. Kostas Plukas, Rimantas Barauskas, Rimvydas Gaidys. Skaitiniai inžinerijos metodai MATLAB'o terpėje. KTU leidykla "Technologija", 2012 m.;
5. R.E. Crochiere and L.R. Rabiner. Multirate Digital Signal Processing. Englewood Cliffs, NJ: Prentice–Hall., 1983 m.;
6. Birne Binegar. Errors in Polynomial Interpolation. Prieiga per internetą: <http://math.okstate.edu/people/binegar/4513-F98/4513-116.pdf>;
7. Fritsch, F. N., R. E. Carlson. Monotone Piecewise Cubic Interpolation. SIAM J. Numerical Analysis, 1980 m.;
8. A. Spitzbart. A Generalization of Hermite's Interpolation Formula. The American Mathematical Monthly, 1960 m.;
9. Pchip. Prieiga per internetą: <http://se.mathworks.com/help/matlab/ref/pchip.html>
10. De Boor, C. A Practical Guide to Splines, Springer-Verlag, 1978 m.;
11. Donatas Levišauskas, Tomas Tekorius. An approach to identification of dynamic model for optimization of fed-Batch fermentation processes. Information, Technology and Control, 2013 m.;
12. D. C. Montgomery. Design and Analysis of Experiments. John Wiley&Sons, 2001 m.;
13. Vytautas Galvanauskas, Donatas Levišauskas. Biotechnologinių procesų modeliavimas, optimizavimas ir valdymas. Praktikumai ir uždaviniai. Vilniaus pedagoginio universiteto leidykla, 2008 m.;
14. V. W. Porto, D. B. Fogel, L. J. Fogel. Alternative neural network training methods. IEEE Expert, 1995 m.;
15. Zhao Chao, Xu Qiaoling, An Aimin, Li Xuelai. Kinetic model parameter estimation by hybrid differential evolution algorithm. Control Conference (CCC), 2012 m.;
16. Donatas Levišauskas, Tomas Tekorius. Model-based optimization of fed-batch fermentation processes using predetermined type feed-rate time profiles. A comparative study. Information, Technology and Control, 2005 m.