RĪGAS TEHNISKĀ UNIVERSITĀTE
Datorzinātnes un informācijas tehnoloģijas fakultāte
Informācijas tehnoloģijas institūts

Artūrs Stepčenko
Doktora studiju programmas "Informācijas tehnoloģija" doktorants

NELINEĀRU UN NESTACIONĀRU NORMALIZĒTA VEĢETĀCIJAS INDEKSA LAIKA RINDU PROGNOZĒŠANAS SISTĒMAS IZSTRĀDĀŠANA

Promocijas darbs

Zinātniskie vadītāji:
profesors *Dr. habil. sc. comp.*
ARKĀDIJS BORISOVS,
profesore *Dr. sc. ing.*
LUDMILA ALEKSEJEVA

Zinātniskais konsultants
vadošais pētnieks *Dr. sc. ing.*
JURIJS ČIŽOVS

RTU Izdevniecība
Rīga 2019

PATEICĪBAS

Izsaku pateicību promocijas darba zinātniskajai vadītājai, profesorei *Dr.sc.ing.* Ludmilai Aleksejevai un zinātniskajam konsultantam, vadošajam pētniekam *Dr.sc.ing.* Jurijam Čižovam par darba tapšanas gaitā sniegto atbalstu, iedrošināšanu, konstruktīvajiem ieteikumiem darba uzlabošanai, kā arī par rekomendācijām un ieteikumiem. Vēlos ar cieņu pieminēt promocijas darba pirmo zinātnisko vadītāju, profesoru *Dr.habil.sc.comp.* Arkādiju Borisovu (1938 – 2016) par atbalsta sniegšanu un iedrošināšanu, sākot promocijas darba izstrādi.

Paldies RTU DITF ITI Modelēšanas un imitācijas katedras akadēmiskajam personālam par palīdzību un ierosinājumiem promocijas darba tapšanas gaitā.

Paldies manai sievai Elīnai par nesavtīgo atbalstu, pacietību un rūpēm par manu profesionālo un personīgo izaugsmi, un par praktisko palīdzību, rediģējot, labojot un formatējot uzrakstītos tekstus. Vēlos pateikties arī meitai Sofijai un dēlam Edgaram par pacietību manā aizņemtajā dzīves posmā.

Paldies mātei Ludmilai, vecvecākiem Lidijai (1937 – 2016) un Dmitrijam, tantei Irinai, sievasmātei Dianai un citiem tuviniekiem un līdzcilvēkiem par morālo atbalstu ilgajā darba tapšanas laikā, kā arī tiešu vai netiešu ieguldījumu darba veikšanā. Izsaku pateicību Ventspils pilsētas domei par finansiālo atbalstu darba tapšanas gaitā.

Šis darbs izstrādāts ar Ventspils pilsētas domes atbalstu saskaņā ar nolikumu par mērķstipendijas piešķiršanu “Atbalsts doktora grāda pretendentiem Ventspils Augstskolā”.

RĪGAS TEHNISKĀ UNIVERSITĀTE
DATORZINĀTNES UN INFORMĀCIJAS TEHNOLOĢIJAS FAKULTĀTE
Informācijas tehnoloģijas institūts

NELINEĀRU UN NESTACIONĀRU NORMALIZĒTA VEĢETĀCIJAS INDEKSA LAIKA RINDU PROGNOZĒŠANAS SISTĒMAS IZSTRĀDĀŠANA

Artūrs Stepčenko

Anotācija

Normalizēta veģetācijas indeksa (*NDVI*) laika rindas ir viens no galvenajiem instrumentiem precīzajā lauksamniecībā. Šī indeksa prognozēšana precīzajā lauksaimniecībā ļauj savlaicīgi noteikt problēmas, kas saistītas ar lauksaimniecības kultūru augšanas tempiem.

Promocijas darbs veltīts nelineāru un nestacionāru *NDVI* indeksa laika rindu analīzei un prognozēšanai ar datu pirmapstrādes, signālu apstrādes, lineāras algebras un mašīnmācīšanās metožu izmantošanu. Promocijas darba mērķis ir izstrādāt normalizēta veģetācijas indeksa laika rindu prognozēšanas sistēmu, kas balstīta uz signālu dekompozīciju un apakšsignāla aproksimācijas pieeju, specializētām datu pirmapstrādes metodēm, kā arī mašīnmācīšanās metodēm.

Promocijas darbā izstrādāta *NDVI* indeksa laika rindu prognozēšanas sistēma (*NDVI FS*), kas veic laika rindas vērtības īstermiņa prognozēšanu nākamajam periodam, izmantojot vēsturiskos novērojumus. Izstrādātās prognozēšanas sistēmas ietvaros realizēta jauna pieeja, kas paredzēta variāciju apakšsignālu dekompozīcijas metodes pielietošanai *NDVI* laika rindu prognozēšanas uzdevumā. Piedāvāta jauna pieeja iegūto datu pirmapstrādes parametru un apmācīta prognozēšanas modeļa pārnesei, kas ļauj šos parametrus un modeli izmantot citu *NDVI* laika rindu prognozēšanai, kur starp apmācībā izmantoto un citu laika rindu ir pietiekami maza Eiklīda distance.

Promocijas darbs ir uzrakstīts latviešu valodā, tajā ir ievads, piecas nodaļas, secinājumi, literatūras saraksts, 36 attēli, 21 tabula, četri pielikumi, kopā 171 lappuse, ieskaitot pielikumus. Literatūras sarakstā ir 121 nosaukums.

RIGA TECHNICAL UNIVERSITY
FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
Institute of Information Technology

FORECASTING SYSTEM DEVELOPMENT FOR NONLINEAR AND NONSTATIONARY TIME SERIES OF NORMALIZED DIFFERENCE VEGETATION INDEX

Artūrs Stepčenko

Annotation

Normalized Difference Vegetation Index (NDVI) time series is one of the most important instruments in precision agriculture. Forecasting of this index in precision agriculture allows us to define problems related to growth rates of agricultural crops in time.

This Doctoral Thesis is devoted to the analysis and forecasting of nonlinear and nonstationary NDVI index time series with the use of data pre-processing, signal processing, linear algebra and machine learning methods. The aim of Doctoral Thesis is to develop a forecasting system of normalized difference vegetation index time series based on signal decomposition and sub-signal approximation approach, specialized data processing methods and machine learning.

In the Thesis, NDVI time series forecasting system (NDVI FS) is proposed that makes short-term time series value forecasting for the next period using historical observations. In the developed system framework, an innovative approach to using variational mode decomposition method in NDVI time series forecasting task is realized. There is also offered an innovative approach to trained forecasting model transfer, which allows using this model to forecast other NDVI time series obtained from neighbourhood pixels.

The Doctoral Thesis has been written in Latvian. It consists of Introduction; 5 chapters; Conclusion; 36 figures; 21 tables; 4 appendices; the total number of pages is 171, including appendices. The Bibliography contains 121 titles.

SATURS

SAĪSINĀJUMU UN APZĪMĒJUMU SARAKSTS 8

IEVADS 10

Tēmas aktualitāte 11

Promocijas darba mērķis un uzdevumi 12

Izvirzītās hipotēzes 12

Pētījuma objekts un priekšmets 13

Pētījuma metodes 13

Darba zinātniskais jaunieguvums un vērtība 14

Darba praktiskais nozīmīgums 14

Darba aprobācija 14

Promocijas darba struktūra 17

1. NORMALIZĒTĀ VEĢETĀCIJAS INDEKSA LAIKA RINDAS UN TO PROGNOZĒŠANAS PĒTĪJUMI 19

1.1. Distancionālās zondēšanas raksturojums 19

1.2. Satelītattēlu izšķirtspējas un to interpretācija 20

1.3. Normalizētā veģetācijas indeksa raksturojums 23

1.4. Pētāmais apgabals un dati 25

1.5. Laika rindas un to īpašības 26

1.6. Zaudējumu funkcijas laika rindu prognozēšanas precizitātes novērtēšanai 28

1.7. Problēmvide un uzdevuma nostādne 30

1.8. Uzdevuma formālā nostādne 31

1.9. *NDVI* laika rindu prognozēšanas pētījumu apskats un analīze 32

1.9.1. Prognozēšanas pētījumu analīze 32

1.9.2. *NDVI* laika rindu prognozēšanas pētījumu kopsavilkums 37

1.10. Laika rindu prognozēšanas pētījumu analīze ar dekompozīcijas metožu pielietojumu 39

1.10.1. Prognozēšanas pētījumu analīze ar dekompozīcijas lietošanu 39

1.10.2. Dekompozīcijas metožu izmantošanas pētījumu kopsavilkums 44

1.11. Prognozēšanas sistēmas konceptuālais apraksts 46

1.12. Kopsavilkums un secinājumi 47

2. PROGNOZĒŠANAS SISTĒMAS IZSTRĀDĀŠANĀ IZMANTOJAMO METOŽU APSKATS 49

2.1. Variāciju apakšsignālu dekompozīcija 49

2.1.1. Furjē transformācija 50

2.1.2. Hilberta transformācija 51

2.1.3. Gausa regularizācija un Vīnera filtrs 51

2.1.4. Variāciju apakšsignālu dekompozīcijas algoritms 52
2.2. Fāžu telpas rekonstrukcija ar laika aiztures metodi 57
2.3. Pakāpeniskā regresijas analīze 58
2.4. Principiālo komponentu analīze 60
2.5. Mākslīgie neironu tīkli 61
2.5.1. Daudzslāņu perceptrons 62
2.5.2. Rekurento slāņu neironu tīkli 66
2.6. Izstrādājamās sistēmas precizitātes salīdzināšanai izmantotās prognozēšanas metodes 67
2.6.1. Vienkāršais slīdošais vidējais 67
2.6.2. Nepārtrauktu stāvokļu kopas Markova ķēdes 68
2.6.3. Automātiski regresējošs integrēts slīdošais vidējais 69
2.7. Kopsavilkums un secinājumi 70

3. APAKŠSIGNĀLA APROKSIMĀCIJAS PIEJAS IZSTRĀDĀŠANA 72
3.1. Oriģinālās *VMD* metodes izmantošana *NDVI* laika rindu prognozēšanai ... 73
3.1.1. Eksperimentu plāns ar oriģinālo *VMD* metodi 73
3.1.2. Oriģinālās *VMD* metodes izmantošana visiem vēsturiskajiem novērojumiem 74
3.1.3. Oriģinālās *VMD* metodes izmantošana katram laika rindas fragmentam 77
3.1.4. Rezultātu analīze eksperimentiem ar oriģinālo *VMD* metodi 79
3.1.5. Oriģinālās *VMD* metodes problēmas *NDVI* laika rindu prognozēšanā 80
3.2. *VMD* metodes modifikācijas izstrādāšana laika rindu prognozēšanai 81
3.3. Eksperimentu plāns ar modificēto *VMD* metodi 83
3.3.1. *NDVI* laika rindu prognozēšana ar modificētās *VMD* metodes izmantošanu 84
3.3.2. *NDVI* laika rindu prognozēšana ar dažādu apakšsignāla precizitāti 86
3.3.3. Rezultātu analīze eksperimentiem ar modificēto *VMD* metodi 87
3.4. Apakšsignāla vērtību tuvinātā aprēķināšana 89
3.4.1. Homogēnās lineāro vienādojumu sistēmas atrisinājums 91
3.4.2. Nehomogēno lineāro vienādojumu sistēmu atrisinājumi 93
3.5. Kopsavilkums un secinājumi 96

4. PROGNOZĒŠANAS SISTĒMAS *NDVI FS* IZSTRĀDĀŠANA 99
4.1. Sistēmas *NDVI FS* funkcijas 99
4.2. Sistēmas *NDVI FS* vispārīgā arhitektūra un darbības princips 100
4.3. Datu pirmapstrādes modulis 103
4.3.1. Apakšsignāla aproksimācija 103

4.3.2. Fāžu telpu rekonstrukcija ... 105
4.3.3. Datu kopu izveidošana ... 106
4.3.4. Pazīmju atlasīšana un izvilkšana ... 108
4.3.5. Pirmapstrādes moduļa skaitļošanas plūsma ... 109
4.4. Mašīnmācīšanās modulis ... 110
4.4.1. Prognozēšanas modeļa apmācība ... 111
4.4.2. *NDVI* laika rindas prognozēšana ... 111
4.5. Izstrādātā prognozēšanas sistēma *NDVI FS* ... 113
4.5.1. Sistēmas *NDVI FS* arhitektūra un skaitļošanas plūsma ... 113
4.5.2. Sistēmas *NDVI FS* praktiskā realizācija ... 115
4.6. Kopsavilkums un secinājumi ... 116

5. IZSTRĀDĀTĀS PROGNOZĒŠANAS SISTĒMAS PRECIZITĀTES NOVĒRTĒŠANA ... 118
5.1. Eksperimentu plāns un apraksts ... 118
5.2. *NDVI* prognozēšana ar vienkāršā slīdošā vidējā izmantošanu ... 121
5.3. Eksperimenta veikšana ar Markova ķēžu izmantošanu ... 122
5.4. *NDVI* prognozēšana ar *ARIMA* ... 123
5.5. Eksperimentu ar sistēmu *NDVI FS* raksturojums ... 124
5.5.1. Prognozēšana bez *VMD* bāzētās aproksimācijas pieejas ... 125
5.5.2. Prognozēšana ar *VMD* apakšsignāla aproksimācijas pieeju ... 129
5.5.3. Prognozēšana ar modeļu pārnešanu bez aproksimācijas pieejas ... 131
5.5.4. Prognozēšana ar modeļu pārnešanu un ar aproksimācijas pieeju ... 137
5.6. Prognozēšanas precizitātes rezultātu apkopojums ... 142
5.7. Prognozēšanas sistēmas *NDVI FS* pielietošana ... 143
5.8. Kopsavilkums un secinājumi ... 146

REZULTĀTI UN SECINĀJUMI ... 148

IZMANTOTĀS LITERATŪRAS SARAKSTS ... 151

PIELIKUMI ... 165
1. pielikums. *NDVI* indeksa pielietošanas veidi mežsaimniecības nozarē ... 166
2. pielikums. *NDVI* indeksa pielietošanas veidi lauksaimniecības nozarē ... 167
3. pielikums. Zaudējumu funkcijas *RMSE* vērtības pieciem eksperimentiem ... 168
4. pielikums. Zaudējumu funkcijas *RMSE* vērtību sadalījums ... 171

SAĪSINĀJUMU UN APZĪMĒJUMU SARAKSTS

Saīsinājums	Nosaukums angļu valodā	Nosaukums latviešu valodā
ADL	*Autoregressive Distributed-Lag Model*	automātiski regresējošs izplatītās aiztures modelis
ADMM	*Alternate Direction Method of Multipliers*	reizinātāju mainīgo virzienu metode
ANN	*Artificial Neural Network*	mākslīgie neironu tīkli
ARIMA	*Autoregressive Integrated Moving Average*	automātiski regresējošs integrēts slīdošais vidējais
ARMA	*Autoregressive Moving Average*	automātiski regresējošs slīdošais vidējais
AVHRR	*Advanced Very High-Resolution Radiometer*	ļoti augstas izšķirtspējas pilnveidots radiometrs
BPFNN	*Back Propagation Feed Forward Neural Network*	kļūdas atgriezeniskās izplatības vienvirziena neironu tīkls
CEEMD	*Complete Ensemble Empirical Mode Decomposition*	pabeigtas grupas empīriskā apakšsignālu dekompozīcija
DS	*Directional Symmetry*	virzienu simetrija
EEMD	*Ensemble Empirical Mode Decomposition*	grupas empīrisko apakšsignālu dekompozīcija
ELM	*Extreme Learning Machine*	ekstrēma apmācības mašīna
EMD	*Empirical Mode Decomposition*	empīriskā apakšsignālu dekompozīcija
ETM+	*The Enhanced Thematic Mapper Plus*	pilnveidotais tematiskās kartografēšanas sensors
EWT	*Empirical Wavelet Transform*	empīriskā veivletu transformācija
GA	*Genetic Algorithm*	ģenētiskais algoritms
LES	*Linear Equations System*	lineāro vienādojumu sistēma
LRNN	*Layer Recurrent Neural Network*	rekurento slāņu neironu tīkls
LSSVM	*Least Squares Support Vector Machine*	mazāko kvadrātu atbalsta vektoru mašīna
MAE	*Mean Absolute Error*	vidējā absolūtā kļūda
MAPE	*Mean Absolute Percentage Error*	vidējā absolūtā procentuālā kļūda
MODIS	*Moderate Resolution Imaging Spectroradiometer*	vidējās izšķirtspējas attēlveidošanas spektroradiometrs
MSE	Mea*n Square Error*	vidējā kvadrātiskā kļūda

Saīsinājums	Nosaukums angļu valodā	Nosaukums latviešu valodā
NDVI	*Normalized Difference Vegetation Index*	normalizētais veģetācijas indekss
NIR	*Reflectance in the Near Infrared Band of the Electromagnetic Spectrum*	atstarošanas koeficients elektromagnētiskā spektra infrasarkanajā joslā
NOAA	*National Oceanographic and Atmospheric Administration*	Nacionālā okeānu un atmosfēras pētījumu pārvalde
OLI	*Operational Land Imager*	operatīvais zemes virsmas attēlu veidotājs
PACF	*Partial Autocorrelation Function*	daļēja autokorelācijas funkcija
PCA	*Principal Component Analysis*	principiālo komponentu analīze
PSO	*Particle Swarm Optimization*	daļiņu spieta optimizācija
PSR	*Phase Space Reconstruction*	fāžu telpas rekonstrukcija
R^2	*Coefficient of Determination*	determinācijas koeficients
RBFNN	*Radial Basis Function Neural Network*	radiālās bāzes funkciju neironu tīkls
RED	*Reflectance in the Red Band of the Electromagnetic Spectrum*	atstarošanas koeficients elektromagnētiskā spektra sarkanajā joslā
R	*Correlation Coefficient*	korelācijas koeficients
RMSE	*Root Mean Square Error*	saknes vidējā kvadrātiskā kļūda
SARIMA	*Seasonal Autoregressive Integrated Moving Average*	sezonālais automātiski regresējošs integrētais slīdošais vidējais
SMA	*Simple Moving Average*	vienkāršais slīdošais vidējais
SNN	*Spiking Neural Network*	impulsu neironu tīkls
STARMA	S*patio-temporal Autoregressive Moving Average*	telpiski-temporālais automātiski regresējošs slīdošais vidējais
SVD	*Singular Value Decomposition*	singulāro vērtību dekompozīcija
TIRS	*Thermal Infrared Sensor*	termālais infrasarkanais sensors
VMD	*Variational Mode Decomposition*	variāciju apakšsignālu dekompozīcija
WD	*Wavelet Decomposition*	veivletu dekompozīcija
WRELM	*Weighted Regularized Extreme Learning Machine*	svērta regularizēta ekstrēma apmācības mašīna

IEVADS

Mūsdienu cilvēku ikdiena ir balstīta uz plānošanu un prognozēšanu. Prognozēšana ir svarīga jebkuras nozares darbībai un attīstībai. Prognozēšanai ir liela nozīme, pieņemot lēmumus lauksaimniecības, ekoloģijas, ekonomikas, finanšu, tehnikas un zinātnes jomās. Tipiski piemēri ir produktu ražības, valūtu un akciju kursu, preču pieprasījuma, meteoroloģiskās u. c. prognozes. Prognoze ir pamatots spriedums par pētāmā objekta iespējamiem stāvokļiem nākotnē, par to sasniegšanas alternatīvajiem ceļiem, nepieciešamajiem resursiem un termiņiem [103]. Regulāra procesu prognozēšana ļauj ne tikai pieņemt efektīvus vadības lēmumus, bet arī uzkrāt pieredzi, kas ļauj paaugstināt prognožu ticamību un precizitāti, kā arī uzlabot prognozēšanas metodes.

Veģetācijas prognozēšana ir cieši saistīta ar daudzām svarīgām starptautiskām problēmām, piemēram, globālo klimata izmaiņu un enerģijas izmantošanas monitoringu, dabas resursu patēriņa pārvaldību, invazīvo augu sugu izplatības prognozēšanu un apdraudēto augu sugu aizsardzību [7]. Datu apkopošana par Zemes virsmas veģetācijas pārklājumu parasti tiek veikta, izmantojot distancionālo zondēšanu. Distancionālā zondēšana ir attālināta Zemes virsmas novērošana ar aviācijas vai satelītu palīdzību, izmantojot dažādus sensorus [73]. Tā kā uz vietas noteikt Zemes virsmas veģetācijas pārklājuma raksturlielumus var tikai nelieliem apgabaliem, tad, izmantojot distancionālo zondēšanu, iespējams veikt Zemes virsmas veģetācijas pārklājuma monitoringu lielā teritorijā, un noteikt veģetācijas izmaiņas, jo šie mērījumi ir gan globāli telpā, gan atkārtoti laikā [68]. Distancionālās zondēšanas dati ir ātri iegūstami un tiek izmantoti jau vairāk nekā 30 gadus, nodrošinot informāciju par ekoloģiskiem procesiem [24]. Distancionālā zondēšana ļauj nomērīt tos apgabalus, kur uz Zemes tiem ir grūti piekļūt [9]. Distancionālajā zondēšanā parasti tiek iegūti satelītattēli, kur katrs attēla pikselis tiek iegūts, konvertējot novērotos zemes virsmas elementus (krāsa, tekstūra, tonis, u. c.) uz skaitlisku vērtību [107].

Laika rindu analīze ir plaši izplatīta metode, lai prognozētu nākotnes izmaiņas apkārtējās vides procesos. Laika rindu analīzei ir divi galvenie mērķi [24]:

1) noteikt fenomenus, kas tiek reprezentēti ar secīgiem novērojumiem, un to raksturīpašības;
2) prognozēt laika rindas nākotnes vērtības.

Abi šie mērķi prasa identificēt un formāli aprakstīt novērotos laika rindas tēlus (*patterns*). Kopš pagājušā gadsimta septiņdesmitajiem gadiem laika rindu prognozēšanai plaši tiek izmantota Boksa–Dženkinsa automātiski regresējoša integrētā slīdošā vidējā (*Autoregressive Integrated Moving Average, ARIMA*) metode [24]. Tomēr *ARIMA* metode, līdzīgi kā lineārās regresijas analīze, ir lineārās laika rindu prognozēšanas metodes, un tas nav robusts pret trokšņiem, kamēr daudzu nozaru praksē izmantotās laika rindas ir nelineāras un trokšņainas [112]. Pēdējās desmitgadēs plaši tiek izmantoti mākslīgie neironu tīkli, un viens no to pielietojumu veidiem ir laika rindu prognozēšana. Atšķirībā no *ARIMA* un lineārās regresijas modeļiem mākslīgie neironu tīkli bieži vien spēj ar labu precizitāti prognozēt nelineāras [62], [114], nestacionāras [14], [47], haotiskas [6], [116] un trokšņainas [15], [20] laika rindas.

Tēmas aktualitāte

Veģetācijas dzīves cikla analīze un prognozēšana ir būtiska, plānojot lauksaimniecības darbus, kā arī veicot lauksaimniecības kultūraugu augšanas tempu monitoringu un prognozējot to ražību. Iegūstot zināšanas par pašreizējām veģetācijas tendencēm un izmantojot šīs zināšanas precīzai prognozēšanai, var līdz minimumam samazināt pārtikas trūkumu.

Bieži praksē tiek izmantoti no satelītattēlu pikseļu vērtībām aprēķinātie veģetācijas indeksi, piemēram, normalizētais veģetācijas indekss (*Normalized Difference Vegetation Index, NDVI*). Dotais indekss apraksta veģetācijas hlorofila daudzumu (zaļo masu), parādot veģetācijas blīvumu un veselību [94]. *NDVI* indekss plaši tiek izmantots ar veģetācijas dinamiku saistītajās nozarēs: lauksaimniecībā, mežsaimniecībā un apkārtējās vides aizsardzībā.

Normalizēta veģetācijas indeksa laika rindas plaši tiek izmantotas precīzajā lauksaimniecībā, kur palīdz pārraudzīt dažādu lauksaimniecības kultūru augšanas tempus un regulēt izlietoto barības vielu apjomu. Šī indeksa prognozēšana precīzajā lauksaimniecībā ļauj savlaicīgi noteikt kultūraugu augšanas problēmas, un laikus pieņemt lēmumus par nepieciešamajiem pasākumiem, lai novērstu šīs problēmas.

Laika rindu prognozēšanas uzdevumos lielu lomu ieņem datu pirmapstrāde [26], [77], [117]. Tiek izmantotas dažādas trokšņu tīrīšanas, nogludināšanas, normalizācijas, transformācijas, pazīmju atlasīšanas un pazīmju izvilkšanas metodes. Tieši šo datu pirmapstrādes metožu un mākslīgo neironu tīklu

kombinācija bieži vien ļauj iegūt labākus prognozēšanas rezultātus [117]. Tādēļ normalizēta veģetācijas indeksa datiem atbilstošu datu pirmapstrādes metožu izvēle, piemērotas prognozēšanas sistēmas un jaunu pieeju izstrādāšana šobrīd ir aktuāls zinātnes jautājums.

Promocijas darba mērķis un uzdevumi

Promocijas darba **mērķis** ir izstrādāt normalizēta veģetācijas indeksa laika rindu prognozēšanas sistēmu, kas balstīta uz signālu dekompozīciju un apakšsignāla aproksimācijas pieeju, specializētām datu pirmapstrādes metodēm, kā arī mašīnmācīšanās metodēm. Promocijas darba mērķa sasniegšanai ir nepieciešams atrisināt tālāk tekstā minētos **uzdevumus**.

1. Normalizēta veģetācijas indeksa laika rindu prognozēšanas metožu un sistēmu salīdzinošs pētījums, lai noteiktu to priekšrocības un iespējamos trūkumus, kas nosaka iegūto *NDVI* laika rindu prognozēšanas precizitāti.
2. Frekvenču analīzei izmantojamo signālu dekompozīcijas metožu salīdzinošs pētījums ar mērķi identificēt metodes, kas prognozēšanas uzdevumos nodrošina augstākas precizitātes sasniegšanu, kā arī noteikt to iespējamās nepilnības.
3. Apakšsignāla aproksimācijas pieejas izstrādāšana normalizēta veģetācijas indeksa laika rindu prognozēšanai.
4. Prognozēšanas sistēmas izstrādāšana, kas izmanto signālu dekompozīcijas metodē balstītu pieeju, datu pirmapstrādes un mašīnmācīšanās metodes.
5. Izstrādātās prognozēšanas sistēmas novērtēšana un precizitātes salīdzināšana ar citām prognozēšanas metodēm.

Izvirzītās hipotēzes

Prognozēšanas sistēmas un apakšsignāla aproksimācijas pieejas izstrādāšanas laikā pārbaudei ir izvirzītas divas hipotēzes.

1. Normalizēta veģetācijas indeksa laika rindu prognozēšanas precizitāte paaugstinās, izmantojot darbā izstrādāto signālu dekompozīcijas metodē balstītu aproksimācijas pieeju.
2. Uz normalizēta veģetācijas indeksa laika rindas apmācītu prognozēšanas modeli un iegūtos datu pirmapstrādes parametrus var izmantot citas līdzīgas *NDVI* laika rindas pirmapstrādei un prognozēšanai ar līdzvērtīgu precizitāti.

Pirmā hipotēze raksturo to, ka vispirms eksperimentāli tiks veikta *NDVI* laika rindu prognozēšana, izmantojot izstrādāto sistēmu, kur datu pirmapstrādes procesā tiks izmantota apakšsignāla aproksimācijas pieeja. Tad tiks veikts eksperiments, prognozējot *NDVI* laika rindas ar izstrādāto sistēmu, bet bez aproksimācijas pieejas izmantošanas. Hipotēze būs uzskatāma par apstiprinātu, ja apakšsignāla aproksimācijas pieejas izmantošana *NDVI* laika rindu prognozēšanas uzdevumā ļaus sasniegt augstāku prognozēšanas precizitāti, nekā bez šīs pieejas izmantošanas.

Otrā hipotēze tiek pamatota idejā, ka iegūtos piemērotos datu pirmapstrādes parametrus un apmācīto prognozēšanas modeli var tālāk izmantot citu *NDVI* laika rindu prognozēšanai. Piemēroti datu pirmapstrādes parametri un prognozēšanas modelis tiek iegūti, uz apmācībā izmantotās *NDVI* laika rindas validācijas datu kopas, iegūstot prognozes ar minimālu saknes vidējo kvadrātisko kļūdu *RMSE*. Ja veicot datu pirmapstrādi un prognozējot kādu no citām *NDVI* laika rindām ar datu pirmapstrādes parametriem un prognozēšanas modeli, kas ir iegūti uz apmācībā izmantotās *NDVI* laika rindas, tiek sasniegta līdzvērtīga prognozēšanas precizitāte, salīdzinot ar to, kas ir iegūta ar citai laika rindai individuāli atrastiem datu pirmapstrādes parametriem un prognozēšanas modeli, hipotēze tiek uzskatīta par apstiprinātu.

Pētījuma objekts un priekšmets

Darba pētījuma **objekts** ir nelineāru un nestacionāru normalizēta veģetācijas indeksa laika rindu prognozēšanas process.

Darba pētījuma **priekšmets** ir datu pirmapstrādes, signālu dekompozīcijas un prognozēšanas metodes, kas ir piemērotas nelineāru un nestacionāru normalizēta veģetācijas indeksa laika rindu prognozēšanai.

Pētījuma metodes

Lai varētu izstrādāt promocijas darbu un apstiprināt izvirzītās hipotēzes, tiek izmantotas vairākas pētījumu metodes. Promocijas darba teorētiskajā izstrādē ir izmantotas šādas metožu klases: mašīnmācīšanās, lineārā algebra, digitālā signālu apstrāde, matemātiskā statistika un varbūtību teorija. Prognozēšanas sistēmas praktiskajai realizācijai tiek izmantota lietojumprogramma *MATLAB*.

Darba zinātniskais jaunieguvums un vērtība

Promocijas darbā ir izstrādāta prognozēšanas sistēma, kura ļauj veikt normalizētā veģetācijas indeksa laika rindu īstermiņa prognozēšanu. Sistēmas izstrādes procesā ir izstrādātas vairākas jaunas pieejas, kā arī metožu un pieeju kopums, kas ir nepieciešami šādas sistēmas realizācijai un novērtēšanai.

1. Izstrādāta no modificētās variāciju apakšsignālu dekompozīcijas metodes iegūta apakšsignāla aproksimācijas pieeja, kas ļauj tuvināti aprēķināt apakšsignāla vērtības katrā laika solī, kur ir pieejami vēsturiskie novērojumi atbilstošai normalizēta veģetācijas indeksa laika rindai.
2. Izstrādāts metožu un pieeju kopums, kas nodrošina augstas precizitātes normalizēta veģetācijas indeksa laika rindu īstermiņa prognozēšanu.
3. Izstrādāta datu pirmapstrādes parametru un prognozēšanas modeļu pārnešanas pieeja, kas nodrošina citu normalizēta veģetācijas indeksa laika rindu prognozēšanu bez jaunas pirmapstrādes un apmācības, ja Eiklīda distance starp apmācībā izmantoto laika rindu un laika rindu, kur izmanto datu pirmapstrādes parametrus un prognozēšanas modeli, ir pietiekami maza.

Darba praktiskais nozīmīgums

Izstrādāta prognozēšanas sistēma, kas ļauj veikt normalizētā veģetācijas indeksa laika rindu īstermiņa prognozēšanu ar labu precizitāti, izmantojot no modificētas variāciju apakšsignālu dekompozīcijas metodes iegūta apakšsignāla aproksimācijas pieeju. Sistēmai ir divi pielietojumu veidi precīzajā lauksaimniecībā.

1. Tā ir vajadzīga, lai varētu savlaicīgi noteikt, kādā apjomā ir nepieciešams izmantot barības vielas (mēslojumu, slāpekli, u.c.), tādējādi nodrošinot, ka nerodas papildu izmaksas.
2. Tā nodrošina iespēju izmantot normalizēta veģetācijas indeksa prognozes kultūraugu ražības modeļos un savlaicīgi aprēķināt sagaidāmos ienākumus.

Darba aprobācija

Promocijas darbā veiktie pētījumi un to rezultāti ir atspoguļoti astoņās publikācijās starptautiskajos, recenzējamos, Latvijas Zinātnes padomes atzītajos zinātniskajos izdevumos.

1. Stepchenko, A. Land Cover Classification Based on MODIS Imagery Data Using Artificial Neural Networks. In: *Proceedings of the 11th International Scientific and Practical Conference* "*Environment. Technology. Resources*", *June 15-17, 2017, Rezekne, Latvia*. Vol. 2. Rezekne: Rezekne Academy of Technologies, 2017, pp. 159-164. ISBN 978-9984-44-173-3. ISSN 1691-5402. e-ISSN 2256-070X. Available from: doi: 10.17770/etr2017vol2.2545. Indexed in: **Scopus**.

2. Stepchenko, A., Chizhov, J., Aleksejeva, L., Tolujew, J. Nonlinear, Non-stationary and Seasonal Time Series Forecasting Using Different Methods Coupled with Data Preprocessing. *Procedia Computer Science*. 2017, vol. 104, pp. 578-585. ISSN 1877-0509. Available from: doi: 10.1016/j.procs.2017.01.175. Indexed in: **Scopus and Web of Science**. Cited: 2.

3. Stepchenko, A., Chizhov, J. Markov Chain Modelling for Short-Term NDVI Time Series Forecasting. *Information Technology and Management Science*. 2016, vol. 19, pp. 39-44. ISSN 2255-9086. e-ISSN 2255-9094. Available from doi: 10.1515/itms-2016-0009. Indexed in: EBSCO, CSA/ProQuest and VINITI.

4. Stepchenko, A. NDVI Index Forecasting using a Layer Recurrent Neural Network Coupled with Stepwise Regression and the PCA. In: *Proceedings of the 5th Virtual International Conference of Informatics and Management Sciences, March 21-25, 2016, Zilina, Slovakia*. Vol. 5, no. 1. Zilina: EDIS-Publishing Institution of the University of Zilina, 2016, pp. 130-135. ISBN 978-80-554-1196-5. ISSN 1339-231X. e-ISSN 1339-9144. Available from: doi: 10.18638/ictic.2016.5.1.281. Indexed in: Google Scholar and Index Copernicus. Cited: 2.

5. Stepchenko, A. Normalized Difference Vegetation Index Forecasting using a Regularized Layer Recurrent Neural Network. In: *Proceedings of the 3rd Virtual Multidisciplinary Conference QUAESTI, December 7-11, 2015, Zilina, Slovakia*. Vol. 3, no. 1. Zilina: EDIS-Publishing Institution of the University of Zilina, 2015, pp. 261-266. ISBN 978-80-554-1170-5. ISSN 2453-7144. e-ISSN 1339-5572. Available from: doi: 10.18638/quaesti.2015.3.1.192. Indexed in: Google Scholar.

6. Stepchenko, A., Chizhov, J. Applying Markov Chains for NDVI Time Series Forecasting of Latvian Regions. *Information Technology and Management Science*. 2015, vol. 18, pp. 57-61. ISSN 2255-9086. e-ISSN: 2255-9094. Available from: doi: 10.1515/itms-2015-0009. Indexed in: EBSCO, CSA/ProQuest and VINITI. Cited: 2.

7. Stepchenko, A., Chizhov, J. NDVI Short-Term Forecasting Using Recurrent Neural Networks. In: *Proceedings of the 10th International Scientific and Practical Conference* "*Environment. Technology. Resources*", *June 18-20, 2015, Rezekne, Latvia*. Vol. 3. Rezekne: Rezeknes Augstskola, 2015, pp. 180-185. ISBN 978-9984-44-173-3. ISSN 1691-5402. e-ISSN 2256-070X. Available from: doi: 10.17770/etr2015vol3.167. Indexed in: **Scopus**. Cited: 2.

8. Stepchenko, A., Borisov, A. Methods of Forecasting Based on Artificial Neural Networks. *Information Technology and Management Science*. 2014, vol. 17, pp. 25-31. ISSN 2255-9086. e-ISSN 2255-9094. Available from: doi: 10.1515/itms-2014-0013. Indexed in: EBSCO, CSA/ProQuest and VINITI.

Promocijas darbā veiktie pētījumi un to galvenie rezultāti ir prezentēti deviņās starptautiskajās zinātniskajās konferencēs.

1. International Conference on Aerospace Engineering, Applied Sciences, Information Technology, Electrical & Mechanical Engineering, Amsterdam, Netherlands, April 27-28, 2019 (Stepchenko, A., Chizhov, J., Aleksejeva, L. Transfer of the Data Preprocessing Parameters and Forecasting Models).

2. 2nd International Conference on Research in Engineering and Fundamental Applied Sciences, Barcelona, Spain, April 20-21, 2019 (Stepchenko, A. An Approach Developing to Approximately Calculate Sub-Signal Obtained from Modified Variational Mode Decomposition Method).

3. Riga Technical University 58th International Scientific Conference, Riga, Latvia, October 12-15, 2017 (Stepchenko, A. Intellectual NDVI Index Forecasting System).

4. 11th International Scientific and Practical Conference "Environment. Technology. Resources", Rezekne, Latvia, June 15-17, 2017 (Stepchenko, A. Land Cover Classification Based on MODIS Imagery Data Using Artificial Neural Networks).

5. Riga Technical University 57th International Scientific Conference, Riga, Latvia, October 14-18, 2016 (Stepchenko, A., Chizhov, J. Short-Term NDVI Index Time Series Forecasting Using Markov Chain with Memory).
6. The 5th International Virtual Scientific Conference on Informatics and Management Sciences, Zilina, Slovakia, March 21-25, 2016 (Stepchenko, A. NDVI Index Forecasting using a Layer Recurrent Neural Network Coupled with Stepwise Regression and the PCA).
7. The 3rd Virtual Multidisciplinary Conference QUAESTI, Zilina, Slovakia, December 7-11, 2015 (Stepchenko, A. Normalized Difference Vegetation Index Forecasting using a Regularized Layer Recurrent Neural Network).
8. Riga Technical University 56th International Scientific Conference, Riga, Latvia, October 14-16, 2015 (Stepchenko, A., Chizhov, J. NDVI Time Series Forecasting Using Markov Chains).
9. 10th International Scientific and Practical Conference "Environment. Technology. Resources", Rezekne, Latvia, June 18-20, 2015 (Stepchenko, A., Chizhov, J. NDVI Short-Term Forecasting Using Recurrent Neural Networks).

Promocijas darba rezultāti ir izstrādāti saistībā ar projektu "Aizaugušu lauksaimniecības platību un neinventarizētu meža zemju meža inventarizācijas rādītāju noteikšana, izmantojot tālizpētes metodes", ko īstenoja SIA "Meža nozares kompetences centrs", Ventspils Augstskola, Latvijas Valsts mežzinātnes institūts "Silava" un SIA "Mikrokods" (14.04.2014 – 30.09.2015).

Promocijas darbs izstrādāts ar Ventspils pilsētas domes atbalstu saskaņā ar nolikumu par mērķstipendijas piešķiršanu "Atbalsts doktora grāda pretendentiem Ventspils Augstskolā".

Promocijas darba struktūra

Promocijas darbā ir ievads, piecas nodaļas, secinājumi, literatūras avotu saraksts un pielikumi. Darbs ir rakstīts latviešu valodā.

Ievadā ir pamatota izvēlētās tēmas aktualitāte, formulēts pētījumu mērķis un uzdevumi, izvirzītas hipotēzes, uzskaitītas promocijas darba izstrādē izmantotās

zinātniskās metodes, aprakstīts pētījumu zinātniskais jaunieguvums un darbā iegūto rezultātu praktiskā vērtība, kā arī sniegts darba aprobācijas raksturojums.

Pirmajā nodaļā aprakstīti normalizēta veģetācijas indeksa teorētiskie pamati, kas iekļauj arī distancionālās zondēšanas un satelītattēlu pirmapstrādes pamatus, aprakstīta esošās situācijas analīze, kas veltīta normalizēta veģetācijas indeksa prognozēšanai un signālu dekompozīcijas metožu izmantošanai laika rindu prognozēšanas uzdevumos.

Otrajā nodaļā aprakstītas datu pirmapstrādes metodes un laika rindu prognozēšanas metodes, kas izmantotas normalizēta veģetācijas indeksa laika rindu prognozēšanas sistēmas *NDVI FS* izstrādāšanā.

Trešajā nodaļā eksperimentāli pārbaudīta oriģinālās variāciju apakšsignālu dekompozīcijas metodes lietojamība normalizēta veģetācijas indeksa laika rindu prognozēšanas uzdevumā. Uz eksperimentu rezultātu analīzes pamata tiek izstrādāta modificēta variāciju apakšsignālu dekompozīcijas metode, kas arī tiek eksperimentāli pārbaudīta. Izstrādāta no modificētās variāciju apakšsignālu dekompozīcijas metodes iegūta apakšsignāla aproksimācijas pieeja.

Ceturtajā nodaļā piedāvāta prognozēšanas sistēma *NDVI FS*. Parādīta sistēmas arhitektūra, ko veido lietotāja saskarne, pirmapstrādes un mašīnmācīšanās moduļi, un datu krātuve. Parādīta moduļu un to veidojošo bloku skaitļošanas plūsma, dots katra bloka apraksts.

Piektajā nodaļā sniegts apraksts eksperimentiem ar izstrādāto sistēmu un iegūtajiem rezultātiem. Novērtēta sistēma ar aproksimācijas pieejas izmantošanu un bez šīs pieejas izmantošanas normalizēta veģetācijas indeksa laika rindu prognozēšanā. Sistēmas prognozēšanas precizitāte salīdzināta ar precizitāti, kas sasniegta, izmantojot klasiskās prognozēšanas metodes. Veikti pirmapstrādes parametru un prognozēšanas modeļu pārnešanas eksperimenti.

Pēdējā nodaļā ietverti darbā sasniegtie rezultāti un secinājumi, kas balstīti uz veiktajiem eksperimentiem un piedāvātas prognozēšanas sistēmas *NDVI FS* izstrādāšanu un lietojumu.

1. NORMALIZĒTĀ VEĢETĀCIJAS INDEKSA LAIKA RINDAS UN TO PROGNOZĒŠANAS PĒTĪJUMI

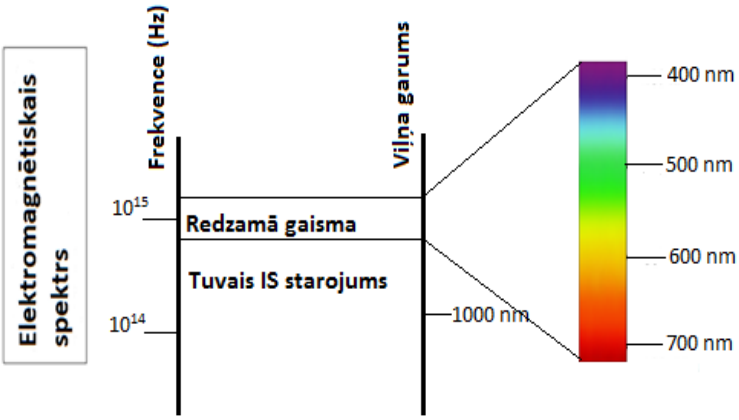
Veģetācijas prognozēšana ir cieši saistīta ar daudzām svarīgām starptautiskām problēmām, piemēram, globālo klimata izmaiņu un enerģijas izmantošanas monitoringu, dabas resursu patēriņa pārvaldību, invazīvo augu sugu izplatības prognozēšanu un apdraudēto augu sugu aizsardzību [7]. Veģetācijas dzīves cikla analīze un prognozēšana ir būtiska, plānojot lauksaimniecības darbus, kā arī veicot lauksaimniecības kultūraugu augšanas un attīstības tempu monitoringu un prognozējot to ražību. Informācijas iegūšana par veģetācijas stāvokli ir svarīgs sākumpunkts, lai analizētu ekosistēmu dinamiku. Datu ievākšana par Zemes virsmas veģetācijas pārklājumu parasti tiek veikta, izmantojot distancionālo zondēšanu.

1.1. Distancionālās zondēšanas raksturojums

Zemes virsmas distancionālā zondēšana (*Remote Sensing*) ir informācijas iegūšana par zemes virsmas objektiem, izmantojot sensorus, kas atrodas attālināti no pētāmā objekta [73]. Distancionālā zondēšana balstās uz fizikālām sakarībām starp reflektēto elektromagnētisko starojumu no zemes virsmas objektiem un šo objektu bioģeofiziskām īpašībām un telpisko izvietojumu [7].

Distancionālā zondēšana tiek veikta, izmantojot sensorus, kuri ir uzstādīti uz satelītiem vai gaisa kuģiem (lidmašīnām, bezpilota lidaparātiem u. c.). Šie sensori reģistrē no zemes virsmas objektiem reflektēto elektromagnētisko starojumu. Elektromagnētiskais starojums ir elektromagnētiskie viļņi, kuri rodas sinusoidāli svārstoties elektriskajam un magnētiskajam laukam kādā telpas punktā [4]. Šīs sinusoidālās svārstības izplatās telpā kā elektromagnētiskais starojums. Elektromagnētiskā starojuma izplatīšanās ātrums ir atkarīgs no vides, vakuumā elektromagnētiskie viļņi izplatās ar gaismas ātrumu $c = 2,99792458x10^8 m/s$, un tos raksturo frekvence $f\left(Hz\right)$ un viļņa garums $\lambda\left(m\right)$. Elektromagnētisko viļņu frekvences var iedalīt skalā, ko sauc par elektromagnētisko spektru [42]. Elektromagnētisko viļņu skalā izšķir vairākus elektromagnētiskā starojuma diapazonus, kas tālāk vēl sadalās joslās, bet promocijas darbā izmantoto vidējās izšķirtspējas attēlveidošanas spektroradiometra normalizētā veģetācijas indeksa (*Normalized Difference Vegetation Index, NDVI*) satelītattēlu ieguvei tiek lietotas redzamās gaismas diapazona sarkanā

josla (*RED*) ar viļņa garumu 620 ÷ 670 nm un tuvā infrasarkanā (*NIR*) starojuma josla ar viļņa garumu 841 ÷ 846 nm (skat. 1.1. att.).



1.1. att. Redzamās gaismas un tuvā infrasarkanā starojuma diapazoni [42].

Redzamo gaismu (optisko spektru) spēj uztvert cilvēka acs. Dažādas redzamās gaismas frekvences cilvēka smadzenes interpretē kā krāsas – no violetās augstākajās frekvencēs, turpinot ar zilu, gaiši zilu, zaļu, dzeltenu un oranžu līdz sarkanajai zemākajās frekvencēs, līdz ar to redzamās gaismas diapazons iedalās vairākās frekvenču joslās.

1.2. Satelītattēlu izšķirtspējas un to interpretācija

Distancionālajā zondēšanā izmantoto sensoru uztvertais elektromagnētiskais starojums tiek saglabāts kādā noteiktā datu formātā, parasti satelītattēlu veidā. Satelītattēli ir rastra digitālie attēli, kas reprezentē novēroto zemes virsmas apgabalu [27]. Digitālais rastra attēls ir matrica, kur katru matricas elementu sauc par attēla pamatelementu jeb pikseli [29]. Katram digitālā attēla pikselim piemīt vērtība, ko sauc par intensitāti un kuru reprezentē ar digitālu skaitli [8]. Pikseli digitālajā attēlā reprezentē atrašanās vietas koordinātas matricā (rindas un kolonnas numurs), bet satelītattēlā to vēl reprezentē arī ģeogrāfiskās koordinātas, kas nosaka pikseļa reālo atrašanas vietu uz zemes virsmas. Satelītattēlus raksturo šādas izšķirtspējas [36].

- **Telpiskā (ģeometriskā) izšķirtspēja**. Raksturo satelītattēla pikselim atbilstošās teritorijas uz zemes virsmas izmēru. Piemēram, telpiskā izšķirtspēja 250 m nozīmē, ka katrs satelītattēla pikselis nosedz 250 m x 250 m lielu teritoriju uz zemes virsmas.

- **Temporālā (laika) izšķirtspēja**. Raksturo satelīta atkārtota apmeklējuma periodu vienam un tam pašam zemes virsmas punktam. Piemēram, temporālā izšķirtspēja septiņas dienas nozīmē, ka satelītam ir vajadzīgas 7 dienas, lai iegūtu jaunu satelītattēlu vienam un tam pašam zemes virsmas punktam.
- **Spektrālā izšķirtspēja**. Atbilst satelītattēla spektrālo (frekvenču) joslu skaitam, kur katra frekvenču josla ir digitāls rastra attēls, kas iegūts kādā noteiktā elektromagnētiskā spektra frekvenču diapazonā.
- **Radiometriskā izšķirtspēja (kontrasts)**. Raksturo, cik daudz informācijas satur viens satelītattēla pikselis, ko izsaka bitos. Pikseļu vērtību diapazonu, ko var pieņemt konkrētā satelītattēla pikseļi, var iegūt, paņemot bitu skaitu bāzē divi pakāpē. Piemēram, ja radiometriskā izšķirtspēja ir 8 biti, tad satelītattēla pikseļu vērtību diapazons pieņem vērtības intervālā [0; 255].

Ja satelītattēls tiek iegūts vairākās elektromagnētiskā starojuma frekvenču joslās, tad šādu attēlu sauc par multispektrālo satelītattēlu [45]. Multispektrālais satelītattēls sastāv no vairākiem digitāliem rastra attēliem jeb joslām, kur šo joslu skaits nosaka satelītattēla spektrālo izšķirtspēju. Ja spektrālo joslu skaits ir mērāms simtos, tad tādu satelītattēlu sauc par hiperspektrālo satelītattēlu [75]. Zemāk redzamajā 1.1. tabulā ir dotas dažu multispektrālo un hiperspektrālo satelītattēlu raksturīpašības un to salīdzinājums. Visi apskatītie satelītattēli tiek iegūti no satelītiem, kuri uz 2017. gada 7. jūliju atradās izplatījumā darba kārtībā.

Vidējās izšķirtspējas attēlveidošanas spektroradiometrs (*Moderate Resolution Imaging Spectroradiometer, MODIS),* ir sensors, ko 1999. gadā palaida izplatījumā uz satelīta *Terra* [67], un tas spēj uztvert datus 36 frekvenču joslās ar telpisko izšķirtspēju 250 m (2 joslas), 500 m (5 joslas) un 1 km (29 joslas). Viens no datu produktiem, ko iegūst no *MODIS Terra* instrumenta, ir normalizēta veģetācijas indeksa (*NDVI*) attēli, kuru telpiskā izšķirtspēja ir 250 m, temporālā izšķirtspēja ir 16 dienas, bet radiometriskā izšķirtspēja ir 16 biti [66]. *MODIS Terra* satelītattēli ir labi piemēroti veģetācijas dinamikas monitoringam, izmantojot veģetācijas indeksu laika rindu analīzi, ilgtermiņa zemes pārseguma izmaiņu monitoringam, kā arī mežu ugunsgrēku izplatības noteikšanai un kartēšanai. Kā redzams 1.1. tabulā, lielākā daļa no šiem satelītattēliem ir pieejami par maksu.

1.1. tabula

Multispektrālo un hiperspektrālo satelītattēlu salīdzinājums

Satelīts un/vai sensors, no kā iegūts attēls	**Telpiskā izšķirtspēja (katrai joslai, metros)**	**Temporālā izšķirtspēja (dienas)**	**Radiometriskā izšķirtspēja (bitos)**	**Cena**
Landsat 7 ETM+ [65]	Panhromatiskā: 15 m; termiskā infrasarkanā: 60 m; pārējās: 30 m	16	8	Bez maksas
Landsat 8 OLI/TIRS [65]	Panhromatiskā: 15 m; termiskā infrasarkanā – 1: 100 m; termiskā infrasarkanā – 2: 100 m; pārējās: 30 m	16	12; Pirmā līmeņa (*Level-1*) datiem: 16	Bez maksas
MODIS Terra [66], [67]	Sarkanā un tuvā infrasarkanā (1–2): 250 m; 3–7 josla: 500 m; pārējās: 1 km	16 (7)	12/16	Bez maksas
MODIS Aqua [66], [67]	Sarkanā un tuvā infrasarkanā (1–2): 250 m; 3–7 josla: 500 m; pārējās: 1 km	16 (7)	12/16	Bez maksas
SPOT-6 [28], [82]	Panhromatiskā: 1,5 m; multispektrālās: 6 m	1	12	Maksas
SPOT-7 [28], [82]	Panhromatiskā: 1,5 m; multispektrālās: 6 m	1	12	Maksas
IKONOS [82]	Panhromatiskā: 0,82 m; multispektrālās: 3,28 m	3	11	Maksas
World-View-2 [82]	Panhromatiskā: 0,46 m; multispektrālās: 1,84 m	1.1	11	Maksas
World-View-3 [82]	Panhromatiskā: 0,31 m; īso viļņu infrasarkanās: 3,7 m; multispektrālās: 1,24 m	1	11	Maksas
World-View-4 [82]	Panhromatiskā: 0,31 m; multispektrālās: 1,24 m	1	11	Maksas

Bez maksas tiek piedāvāti *Landsat 7* pilnveidota tematiskās kartografēšanas sensora (*The Enhanced Thematic Mapper Plus*, *ETM+*), *Landsat 8* operatīvā zemes virsmas attēlu veidotāja/termāla infrasarkanā sensora (*Operational Land Imager/ Thermal Infrared Sensor*, *OLI/TIRS*), *MODIS Terra* un *Aqua* satelītattēli. Promocijas darbā tiek izmantoti *MODIS Terra* satelītattēli, pateicoties to trim priekšrocībām.

- Ir pieejamas tiešsaistes platformas, kuras nodrošina jau gatavus *NDVI* attēlus, kas iegūti no *MODIS Terra* satelītattēliem, ar jau veiktu attēlu pirmapstrādi.
- *MODIS Terra NDVI* attēliem ir pieejama temporālā izšķirtspēja 7 vai 16 dienas. Tātad augstākā pieejamā temporālā izšķirtspēja ir 7 dienas, kamēr *Landsat* satelītattēliem tā ir 16 dienas;
- Tā kā satelīts *Terra* ir palaists izplatījumā par aptuveni 3 gadiem agrāk nekā satelīts *Aqua*, tad *MODIS Terra* satelītattēliem ir pieejams lielāks attēlu skaits, līdz ar to ir iespējams iegūt laika rindas ar lielāku novērojumu skaitu, nekā no *MODIS Aqua* satelītattēliem.

Septiņu dienu periods ir piemērots veģetācijas attīstības dinamikas monitoringam (novērošanai) un prognozēšanai precīzajā lauksaimniecībā.

1.3. Normalizētā veģetācijas indeksa raksturojums

Normalizētais veģetācijas indekss (*NDVI*) ir skaitlisks fotosintētiski aktīvas biomasas rādītājs [108], kas tiek aprēķināts, izmantojot atstarošanas koeficienta vērtības divās elektromagnētiskā spektra joslās pēc izteiksmes (1.1.):

$$NDVI = \frac{NIR - RED}{NIR + RED}, \qquad (1.1.)$$

kur NIR – atstarošanas koeficienta vērtība tuvajā infrasarkanajā joslā;

RED – atstarošanas koeficienta vērtība redzamās gaismas sarkanajā joslā.

Šīm divām joslām piemīt šādas īpašības: sarkanajā spektrā ir sakoncentrēta maksimālā Saules radiācijas absorbcija ar hlorofila palīdzību, bet infrasarkanajā joslā atrodas maksimālais veģetācijas zaļo sastāvdaļu refleksijas apgabals [78]. Lai attēlotu *NDVI* indeksa vērtības, tiek izmantota standarta gradienta skala, kur indekss pieņem vērtības no [–1; 1] vai diskrēta skala, mērogota intervālā [0; 255], kas atbilst pelēko toņu gradāciju līmeņiem.

Normalizētais veģetācijas indekss ir mēreni jutīgs pret augsnes un atmosfēras slāņu izmaiņām, bet izteikti reaģē uz veģetācijas slāņa izmaiņām [120]. Pateicoties atstarošanas īpatnībām redzamās gaismas sarkanajā un tuvajā infrasarkanajā elektromagnētiskā spektra joslā, dabas objekti, kuri nav saistīti ar veģetāciju, pieņem fiksētas *NDVI* vērtības (skat. 1.2. tabulu).

1.2. tabula

Objekti un to *NDVI* indeksa vērtības [120]

Objekta tips	**Refleksijas vērtība sarkanajā spektrā (%)**	**Refleksijas vērtība infrasarkanajā spektrā (%)**	***NDVI* indeksa vērtība**
Bieza veģetācija	0,1	0,5	0,7
Vāja veģetācija	0,1	0,3	0,5
Atklāta augsne	0,25	0,3	0,025
Mākoņi	0,25	0,25	0
Sniegs un ledus	0,375	0,35	–0,05
Ūdens	0,02	0,01	–0,25
Mākslīgi materiāli (betons, asfalts)	0,3	0,1	–0,5

Negatīvās *NDVI* vērtības norāda uz mākslīgiem materiāliem, piemēram, betonu un asfaltu, kā arī uz ūdens virsmu. Vērtības ap nulli norāda uz klinšainu vai sniegotu virsmu. Nulles vērtība raksturo mākoņu pārklājumu. Zemās un vidējās *NDVI* vērtības norāda uz krūmājiem un pļavām, bet augstās *NDVI* vērtības norāda uz blīvu veģetācijas virsmu, piemēram, mežiem [69].

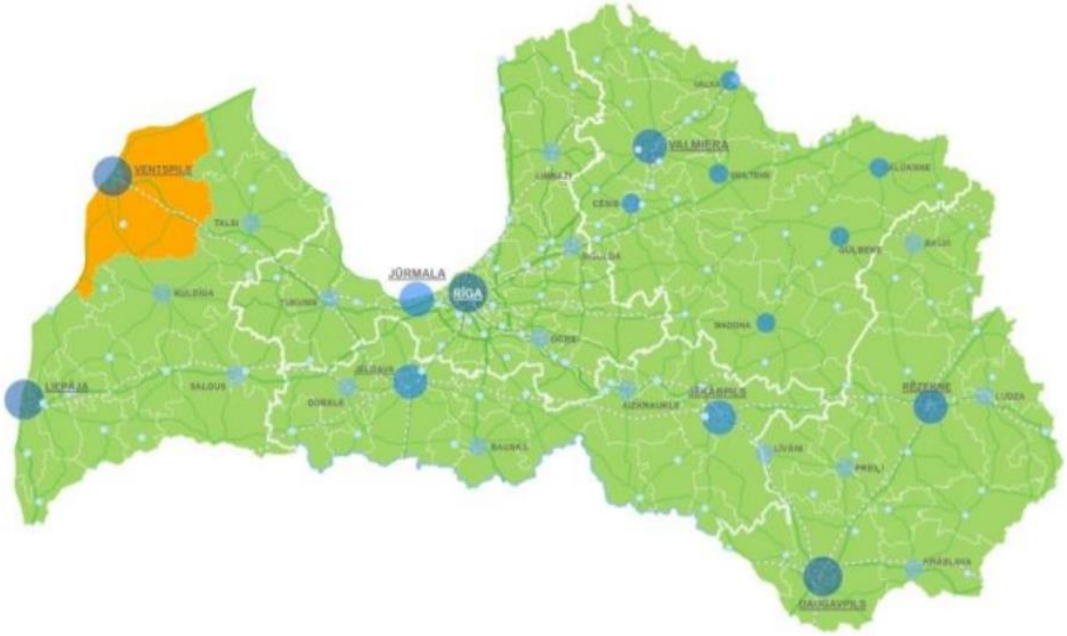
Pastāv cieša korelācija starp *NDVI* indeksu un dažādu ekosistēmu (piemēram, lauksaimniecības zemju, skujkoku mežu, lapu koku mežu, jaukto mežu) produktivitāti. Šī īpašība tiek izmantota, lai kartētu un analizētu dažādas ekosistēmas, novērtētu biosistēmu resursus un platības. Normalizēto veģetācijas indeksu galvenokārt izmanto lauksaimniecības un mežsaimniecības nozarēs. Galvenie *NDVI* indeksa pielietošanas veidi mežsaimniecības nozarē ir reģiona bioresursu un mežu produktivitātes un mežsaimniecības efektivitātes novērtēšana, mežu ugunsgrēku izplatības, mežu izciršanas un mežu robežu izmaiņu monitorings un analīze, un nekoksnes mežu resursu ražības novērtēšana, un šo pielietošanas veidu apraksti un tajos veicamie uzdevumi ir doti **1**. pielikumā. Savukārt *NDVI* indeksa pamata pielietošanas veidi lauksaimniecības nozarē ir kultūraugu sējumu kartēšana, kultūraugu attīstības tempu monitorings un analīze, un ražības prognozēšana, un šo pielietošanas veidu apraksti un tajos veicamie uzdevumi ir apskatīti **2**. pielikumā.

Lai gan *NDVI* indekss ir mākslīgs rādītājs, kas paredzēts ekoloģisku un klimatisku izmaiņu mērīšanai, tam pastāv arī ievērojama korelācija ar citu sfēru

parametriem: produktivitāti, biomasas apjomu, mitrumu un augsnes minerālu piesātinājumu, iztvaikošanu, nokrišņu apjomu, sniega pārklājuma raksturlielumiem [120]. Sakarība starp šiem parametriem un *NDVI* indeksu gan nav tieša, tādēļ *NDVI* indeksu bieži izmanto kā starpposmu, lai veiktu sarežģītāku analīzi.

1.4. Pētāmais apgabals un dati

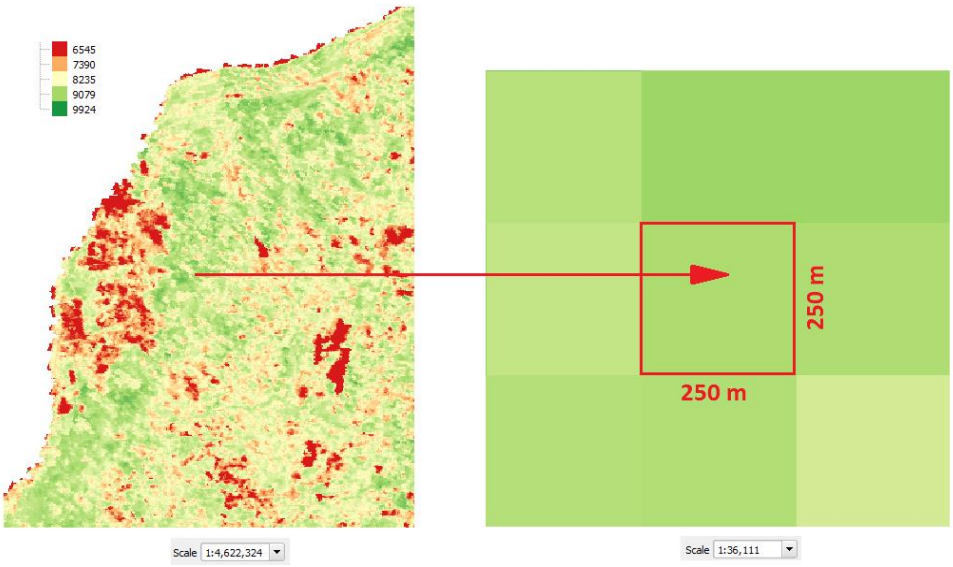
Darbā izmantotais pētījumu apgabals ir Ventspils novads, kas atrodas Latvijas ziemeļrietumu daļā un aizņem 2457,56 km^2 lielu platību [104]. Ventspils novads Latvijas kartē parādīts 1.2. attēlā, atzīmēts ar oranžu krāsu. Lauksaimniecībā izmantojamā zeme Ventspils novadā aizņem 511 000 ha lielu platību jeb 20,9 % no novada teritorijas platības, savukārt mežu zemes aizņem 64 %, no kopējās novada teritorijas [104].



1.2. att. Ventspils novads Latvijas kartē [104].

Promocijas darba datu kopu veido 814 nogludināti *NDVI* attēli, kas iegūti no *MODIS Terra* satelītattēliem ar telpisko izšķirtspēju 250 m, temporālo izšķirtspēju septiņas dienas un radiometrisko izšķirtspēju 16 biti. Šie attēli pārklāj Ventspils novada teritoriju. Pikseļu vērtības atrodas intervālā [–10000; 10000], kuru, reizinot ar konstanti 0,0001, tiek iegūts standarta *NDVI* vērtību diapazons [–1; 1]. *NDVI* attēli tika lejupielādēti no Vīnes Dabas resursu un dabaszinātņu universitātes *MODIS* veģetācijas indeksu laika rindu datu apstrādes platformas, uz kuras arī tiek veikta *MODIS Terra*

NDVI attēlu pirmapstrāde, kas iekļauj pikseļu vērtību nogludināšanu un trūkstošo vērtību aizpildīšanu, kas radušās mākoņu pārklājuma vai sliktu atmosfēras nosacījumu dēļ [106]. Zemāk redzamajā 1.3. attēlā dots piemērs vienam šādam *MODIS Terra* satelītattēlam, kas iegūts 2016. gada 19. jūlijā.



1.3. att. Ventspils novada *MODIS* satelītattēls.

NDVI attēls kreisajā pusē ģeogrāfiski aptver Ventspils novadu, bet labajā pusē ir dota šī paša attēla detalizācija līdz vienam pikselim – 250 m x 250 m lielam apgabalam. *NDVI* attēla mazākais pamatelements ir pikselis, kura vērtība arī ir *NDVI* indeksa vērtība, kas raksturo zemes virsmas pārklājumu šajā apgabalā, kas *MODIS Terra NDVI* attēliem ir 250 m x 250 m liels apgabals. Izvietojot *NDVI* attēlus hronoloģiskā secībā, katram pikselim tiek iegūta secīgu *NDVI* vērtību virkne, kas ļauj iegūt skalāru *NDVI* laika rindu.

1.5. Laika rindas un to īpašības

Laika rinda ir novērojumu (mērījumu) kopa, kas ir sakārtota hronoloģiskā secībā [1]. Parasti laika rindas novērojumi tiek veikti starp vienādiem laika intervāliem, tomēr pastāv arī laika rindas, kur laika intervāli starp novērojumiem nav vienādi. Matemātiski laika rinda y tiek definēta kā vektors, kuru apraksta izteiksme (1.2.):

$$y = \{y(t), t = 1, 2, 3, \ldots, N\}, \tag{1.2.}$$

kur $y(t)$ – laika rindas y vērtība laika solī t;

t – laika solis;

N – novērojumu skaits.

Ja laika rinda satur viena mainīgā novērojumus, tad tādu laika rindu sauc par vienfaktora (*univariate*) laika rindu, bet, ja tā satur vairāku mainīgo novērojumus, tad tādu laika rindu sauc par daudzfaktoru (*multivariate*) laika rindu [1].

Laika rindas parasti satur četras galvenās komponentes, kas var tikt izdalītas novērotajos datos [1]:

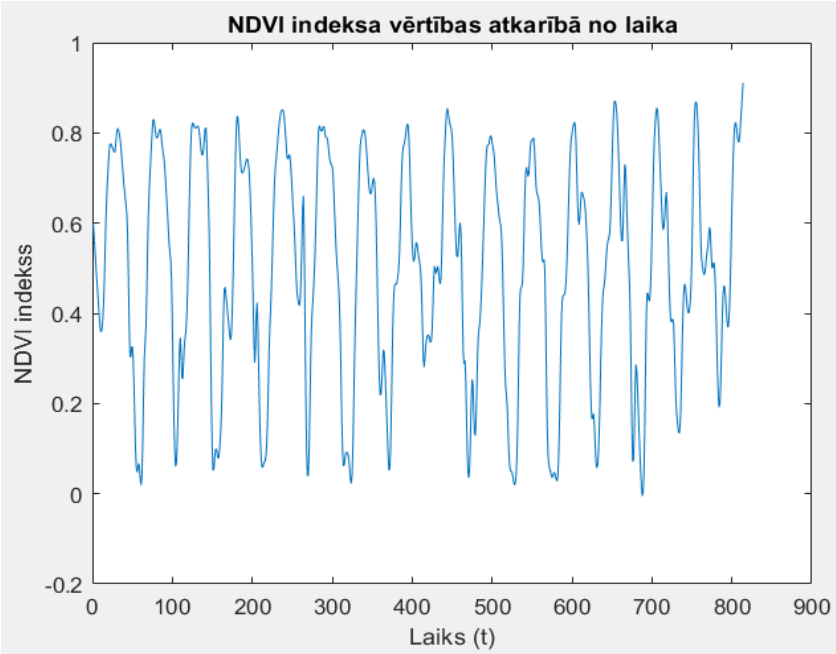
- trends – laika rindas tendence pieaugt, krist vai stāvēt uz vietas ilgtermiņā; var teikt, ka trends ir laika rindas ilgtermiņa kustība;
- cikliskā komponente – laika rindas svārstības vidējā termiņā, kuras izraisa noteikti apstākļi, kas atkārtojas cikliski;
- sezonālā komponente – laika rindā novērojamās svārstības gada laikā atbilstoši sezonām, ko izraisa tādi faktori kā klimats un laika apstākļi;
- neregulārā komponente – neregulārās jeb gadījuma izmaiņas laika rindā, kuri rodas no neprognozējamiem neregulāriem faktoriem un neatkārtojas laika rindas tēlos (*patterns*).

Nelineārā laika rinda $y(t)$ ir signāls, kas tiek iegūts no nelineāra, dinamiska procesa [21]. Šis signāls ir partikulārs risinājums nelineāram, stohastiskam diferenciālvienādojumam, kas apraksta kādu sistēmu. Tādējādi daudzu nelineāru, stohastisku un dinamisku sistēmu risinājums ir nelineāra laika rinda.

Stacionārā laika rinda ir tāda laika rinda, kuras statistiskās īpašības, piemēram, vidējā vērtība, dispersija, autokorelācija u.c. ir konstantas laika gaitā [109]. Tomēr daudzas nelineārās un dinamiskās sistēmas darbojas mainīgos (t. i., nestacionāros) apstākļos. Jo lielāks ir laika posms, kurā ir iegūti laika rindas vēsturiskie novērojumi, jo lielāka ir iespēja, ka laika rinda būs nestacionāra. Parasti laika rindas, kas satur trendu vai sezonālās svārstības, arī ir nestacionāras laika rindas [1].

Šajā promocijas darbā izmantotās normalizēta veģetācijas indeksa (*NDVI*) laika rindas ir nelineāras un nestacionāras vienfaktora laika rindas ar aditīva trokšņa un sezonālās komponentes elementiem [37]. Katra laika rindas vērtība ir *NDVI* indeksa vērtība, kas iegūta noteiktā datumā, un katrs laika rindas elements iegūts ar vienas nedēļas periodu.

Visu *NDVI* laika rindas novērojumu skaits $N = 814$. Vienas normalizēta veģetācijas indeksa laika rindas piemērs dots 1.4. attēlā.



1.4. att. *NDVI* indeksa laika rinda.

No 1.4. attēla ir redzams, ka *NDVI* laika rindai piemīt izteikta sezonāla komponente. Maksimālās vērtības *NDVI* laika rinda 1.4. attēlā sasniedz vasaras periodā, kad augos ir sastopams maksimālais hlorofila daudzums. Laika rindas maksimālā vērtība ir 0,91. Savukārt zemākās vērtības tiek sasniegtas ziemas periodā ar minimālo vērtību –0,005, bet laika rindas vidējā vērtība ir 0,4965. Tāpat laika rindas datos ir redzama trokšņa komponente, kas apgrūtina lineāru prognozēšanas modeļu izmantošanu gan šīs, gan citu *NDVI* laika rindu prognozēšanai.

1.6. Zaudējumu funkcijas laika rindu prognozēšanas precizitātes novērtēšanai

Saknes vidējā kvadrātiskā kļūda (*Root Mean Squared Error*, *RMSE*) parāda kļūdu kvadrātu vidējās vērtības sakni starp novērotās un prognozētās laika rindas vērtībām katrā laika solī [110], un to var izteikt ar izteiksmi (1.3.):

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}\left(\hat{y}(t) - y(t)\right)^2}{N}}\ , \qquad (1.3.)$$

kur $\hat{y}(t)$ – prognozēta laika rinda solī t;

$y(t)$ – novērota laika rinda solī t;

N – novērojumu skaits.

Virzienu simetrija (*Directional Symmetry*, *DS*) parāda, cik procentuāli ir pareizi prognozētās laika rindas virzienu izmaiņas, no viena laika soļa uz nākamo [110]. Virzienu simetriju *DS* var izteikt ar izteiksmi (1.4.):

$$DS = \frac{100}{N-1}\sum_{t=2}^{N} d_t \,, \tag{1.4.}$$

kur $d_t = \begin{cases} 1, \text{ ja } \left(\hat{y}(t) - y(t-1)\right)\left(y(t) - y(t-1)\right) \geq 0; \\ 0, \text{ pretējā gadījumā.} \end{cases}$

Determinācijas koeficients R^2, ko viena atribūta gadījumā aprēķina kā korelācijas koeficienta R kvadrātu, ir prognozējama lieluma (šajā darbā – laika rindas) dispersijas proporcija, ko var izskaidrot, izmantojot neatkarīgos parametrus jeb atribūtus [32]. Determinācijas koeficientu R^2 var izteikt ar izteiksmi (1.5.):

$$R^2 = 1 - \frac{\sum_{t=1}^{N}\left(\hat{y}(t) - y(t)\right)^2}{\sum_{t=1}^{N}\left(y(t) - \frac{1}{N}\sum_{t=1}^{N} y(t)\right)^2} \,. \tag{1.5.}$$

Ja atribūtu skaits prognozēšanas modelī ir lielāks par viens, tad izmanto pielāgoto R^2 [72], kuru apraksta izteiksme (1.6.):

$$R_{adj}^2 = 1 - \frac{(N-1)}{(N-p)}(1 - R^2) \,, \tag{1.6.}$$

kur p – atribūtu skaits prognozēšanas modelī.

Ja prognozēšanas modelī ir vairāk nekā viens atribūts (neatkarīgais parametrs), tad, pieaugot atribūtu skaitam, pieaug arī determinācijas koeficienta vērtība. Lai novērstu vērtības pieaugumu, tiek izmantots pielāgotais R_{adj}^2, kur vērtības pieaugums nav proporcionāls atribūtu skaitam.

1.7. Problēmvide un uzdevuma nostādne

Precīzā lauksaimniecība ir uz plašu, vispusīgu, precīzu un savlaicīgu informāciju balstīta saimniekošana laukos, sākot ar saimniekošanas stratēģijas noteikšanu, ražošanas tehnoloģiju un tehniskā nodrošinājuma plānošanu, produkcijas ražošanu, pēcapstrādi, uzglabāšanu un beidzot tās realizācija [105]. Piemēram, kviešu attīstības monitorings ļauj koriģēt olbaltumvielu deficītu, izmantojot slāpekli līdz pat augšanas sezonas beigām [83].

Satelītattēli un no tiem aprēķinātais normalizētais veģetācijas indekss (*NDVI*), kas, piemēram, tiek iegūts ar nedēļas intervālu, nodrošina iespēju salīdzināt pašreizējā gada kviešu pieaugumu konkrētajā nedēļā ar iepriekšējo gadu kviešu pieaugumu šajā pašā nedēļā. Par labu precīzās lauksaimniecības izmantošanai var tikt minēti četri aspekti: ražas apjoma kāpināšana, materiālo ieguldījumu samazināšana, ekonomiskais izdevīgums un ekoloģiskie apsvērumi [105].

1. Ražas apjoma kāpināšana. Daudzi lauksaimnieki izmanto lauku diferencētu mēslošanu. Tas nozīmē, ka mēslojums netiek vienmērīgi izlietots pa visu lauku, bet pēc noteikta plāna – tas tiek palielināts vai samazināts dažādos lauka nogabalos.
2. Materiālo ieguldījumu samazināšana. Zinot ražu samazinošos faktorus konkrētā lauka ietvaros, ir iespēja aprēķināt šī lauka ražas kāpināšanas optimālo risinājumu. Tādā veidā ir iespējams ekonomēt mēslojumu, herbicīdus, degvielu u. c. materiālos resursus.
3. Ekonomiskais izdevīgums. Kultūraugu raža var tikt iegūta ar mazāku līdzekļu ieguldījumu. Var tikt sastādītas prognozējamo un gūto ienākumu kartes pa lauka nogabaliem.
4. Ekoloģiskie apsvērumi. Ja mēslojums un pesticīdi tiek izmantoti tikai konkrētās lauka vietās iespējami minimālos apjomos, tas samazina lauksaimnieciskās darbības nevēlamo ietekmi uz apkārtējo vidi.

Tātad *NDVI* indekss ļauj sekot līdzi kultūraugu attīstībai, salīdzinot indeksa vērtības konkrētos datumos ar iepriekšējo gadu vērtībām šajos pašos datumos. Otrs indeksa pielietojums precīzajā lauksaimniecībā ir ražības un sagaidāmo ienākumu aprēķināšana, kur *NDVI* tiek izmantots kā parametrs dažādos vienādojumos.

NDVI indeksa īstermiņa prognozēšana vienu nedēļu uz priekšu nodrošina iespēju salīdzināt nevis pašreizējo situāciju ar kultūraugu attīstību, bet prognozējamo.

Tas ļauj savlaicīgi pieņemt nepieciešamos lēmumus, piemēram, par mēslojuma palielināšanu vai samazināšanu.

Promocijas darbā tiek risināts prognozēšanas uzdevums, kur ir jāaprēķina *NDVI* laika rindas īstermiņa prognoze vienu nedēļu uz priekšu, izmantojot datu pirmapstrādes, signālu dekompozīcijas, lineārās algebras un mašīnmācīšanās metodes.

1.8. Uzdevuma formālā nostādne

Ieejas laika rindas $y = \{y(t)\}$ novērojumu vērtības ir pieejamas laika momentos $t = 1, \ldots, N$. No laika rindas y ir iegūtas atvasinātas vērtības, kas arī ir laika rindas $u_1 = \{u_1(t)\}, u_2 = \{u_2(t)\}, \ldots, u_K = \{u_K(t)\}$, kas tāpat ir pieejamas laika momentos $t = 1, \ldots, N$. Laika brīdī $t = N$ ir nepieciešams noteikt prognozi ieejas laika rindas $y = \{y(t)\}$ vērtībai laika brīdī $t = N + 1$, ņemot vērā arī atvasināto laika rindu vērtības.

Lai noteiktu procesa $y = \{y(t)\}$ novērojuma vērtību laika brīdī $t = N + 1$, ir nepieciešams noteikt funkcionālu sakarību starp $y = \{y(t)\}$ vēsturiskajām novērojumu vērtībām un nākotnes vērtībām, kā arī šai sakarībai ir jāņem vērā atvasināto laika rindu vēsturisko vērtību ietekmi uz ieejas laika rindu kā to apraksta (1.7.) vienādojums:

$$\begin{aligned} y(t+1) = f(&y(t), y(t-1), \ldots, y(t-m+1), u_1(t), u_1(t-1), \ldots, \\ &u_1(t-m_1+1), \ldots, u_K(t), u_K(t-1), \ldots, u_K(t-m_K+1)) + \varepsilon(t), \end{aligned} \tag{1.7.}$$

kur y – ieejas laika rinda;

K – atvasināto laika rindu skaits;

u_1 – pirmā atvasinātā laika rinda;

u_K – *K*-tā atvasinātā laika rinda;

m – loga garums, ko izmanto ieejas laika rindai;

m_1 – loga garums, ko izmanto pirmajai atvasinātajai laika rindai;

m_K – loga garums, ko izmanto *K*-tai atvasinātajai laika rindai;

t – laiks;

ε – baltais troksnis.

Funkcionālā sakarība (1.7. vienādojums) ir prognozēšanas modelis. Ir nepieciešams iegūt tādu prognozēšanas modeli, kuram zaudējumu funkcijas – saknes vidējā kvadrātiskā kļūda *RMSE* starp patiesajām un prognozētajām laika rindas $y = \{y(t)\}$ vērtībām ir minimāla.

1.9. *NDVI* laika rindu prognozēšanas pētījumu apskats un analīze

Normalizētā veģetācijas indeksa (*NDVI*) laika rindu īstermiņa prognozēšana ir vitāli svarīga, pieņemot lēmumus precīzajā lauksaimniecībā. Zinātniskajā literatūrā *NDVI* indeksa laika rindu prognozēšana bieži tiek veikta ar autoregresijas metodēm, bet tiek izmantoti arī vienvirziena mākslīgie neironu tīkli un lineārā regresijas analīze. Daži *NDVI* indeksa prognozēšanas pētījumi ir apkopoti 1.3. tabulā.

1.9.1. Prognozēšanas pētījumu analīze

Pētījumā [24] tiek veikta *NDVI* indeksa laika rindas īstermiņa prognozēšana, izmantojot Boksa–Dženkinsa metodes: automātiski regresējošo integrēto slīdoša vidējā (*ARIMA*) modeli un sezonālo *ARIMA* modeli (*SARIMA*). Prognoze tiek veikta 10 dienu uz priekšu. Izmantotās *NDVI* laika rindas tiek iegūtas no ģeometriski un radiometriski koriģētiem Nacionālās okeānu un atmosfēras pētījumu pārvaldes ļoti augstas izšķirtspējas pilnveidotā radiometra (*National Oceanographic and Atmospheric Administration Advanced Very High-Resolution Radiometer*, *NOAA AVHRR*) multispektrāliem satelītattēliem, un pētāmais apgabals ietver Spānijas ziemeļu centrālo daļu. Kā papildu dati tiek izmantota koku sugu digitālā karte, kas ļauj iegūt *NDVI* attēlus atbilstoši katrai koku sugai. Kopā tiek iegūtas sešas *NDVI* attēlu laika rindas ar 185 attēliem katrā no tām. No katra *NDVI* attēla tiek paņemta visu attēlā esošo pikseļu vērtību vidējā vērtība, un tiek iegūtas sešas skalāras *NDVI* laika rindas, atbilstoši katrai koku sugai. Kā labākais prognozēšanas modelis tiek atrasts sezonālais *ARIMA* modelis (1.8. izteiksme):

$$SARIMA(p,d,q)\,x\,(P,D,Q)_S, \qquad (1.8.)$$

kur p – autoregresīvā polinoma kārta, $p = 0$;

d – diferenču operatora (integrācijas) kārta, $d = 1$;

q – slīdošā vidējā polinoma kārta, $q = 1$;

P – sezonālā autoregresīvā polinoma kārta, $P = 0$;

D – sezonālā diferenču operatora kārta, $D = 1$;

Q – sezonālā slīdošā vidējā polinoma kārta, Q=1;

S – sezonālās komponentes periods, S=37.

1.3. tabula

NDVI indeksa pētījumu apskats ārzemju zinātniskajā literatūrā

Prognozēšanas metode	**Papildu ieejas dati**	**Kļūdu un precizitātes rādītāji**			
ARIMA modelis un *SARIMA* modelis [24]	Nav	$RMSE = 0,0007 \div 0,0015$.	$MAE = 0,02 \div 0,03$.	$MAPE = 9,15 \div 10,88\,\%$.	
ARIMA modelis [5]	Nav	$R = 0,62$ un $R = 0,077$.			
STARMA modelis un *ARIMA* modelis [13]	Nav	$MSE = 1,96 \div 2,10$ *STARMA* modelim.		$MSE = 2,07 \div 2,25$ *ARIMA* modelim.	
ANN un Naivais Beijesa prognozēšanas modelis [61]	Nav	*ANN* modeļa precizitāte ir augstāka nekā Naivā Beijesa prognozēšanas modeļa precizitāte.			
ADL modelis [40]	Nokrišņu un temperatūras dati	2 nedēļu prognoze	4 nedēļu prognoze	8 nedēļu prognoze	12 nedēļu prognoze
		$R^2 = 0,96$; $MAPE = 5,0\,\%$.	$R^2 = 0,93$; $MAPE = 7,4\,\%$.	$R^2 = 0,85$; $MAPE = 11,1\,\%$.	$R^2 = 0,79$; $MAPE = 12,9\,\%$.
Vienvirziena neironu tīkls [60]	Dienvidu svārstību indekss	$R^2 = 0,79$; $RMSE = 0,01$.			
Daudzparametru lineārā regresijas analīze [37]	Augsnes mitruma dati, zemes izmantojuma dati un jūras virsmas temperatūras dati	Labākā precizitāte ir daudzparametru lineārās regresijas modelim ar trim ieejas parametriem (augsnes mitruma, jūras virsmas temperatūras un *NDVI* indeksa vēsturiskajiem datiem).			

Saknes vidējā kvadrātiskā kļūda (*RMSE*) uz sešu prognozēto laika rindu validācijas datu kopām svārstās robežās starp vērtībām 0,0007 līdz 0,0015, bet vidējā absolūtā kļūda (*Mean Absolute Error, MAE)* svārstās robežās starp 0,02 līdz 0,003. Savukārt vidējā absolūtā procentuālā kļūda (*Mean Absolute Percentage Error, MAPE*) svārstās robežās starp 9,15 % līdz 10,88 %.

Īstermiņa *NDVI* laika rindas prognozēšana ar prognozēšanas periodu viens mēnesis tiek veikta pētījumā [5]. Pētījumu apgabals ietver divus reģionus Dienvidamerikā, Amazones baseinā. Kā prognozēšanas modelis tiek izmantots *ARIMA* modelis. No abu reģionu *NOAA AVHRR NDVI* attēlu laika rindām tiek iegūtas divas skalāras *NDVI* laika rindas, ņemot katra attēla pikseļu vērtību vidējo vērtību. Prognozējot pirmo *NDVI* laika rindu, tiek iegūts korelācijas koeficients R =0,6204 starp prognozēto un novēroto laika rindu, savukārt prognozējot otro *NDVI* laika rindu, tiek iegūts ļoti zems korelācijas koeficients R =0,07746.

Lai prognozētu *NDVI* indeksa laika rindu 16 dienas uz priekšu, tiek izmantota telpiski-temporāla automātiski regresējoša slīdošā vidējā (S*patio-temporal Autoregressive Moving Average, STARMA*) metode [13]. *STARMA* metode ir lineāra metode, kas izmanto ne tikai kāda pikseļa vēsturiskās vērtības, bet, ņemot vērā kaimiņu pikseļu korelāciju, arī šī pikseļa kaimiņu pikseļu vērtības. Pētāmais apgabals ietver dažus Austrumāfrikas reģionus. Piecas *NDVI* laika rindas pikseļu līmenī tiek iegūtas no ģeometriski un radiometriski koriģētiem *NOAA AVHRR* multispektrālajiem attēliem, un tās katra satur 600 novērojumus. Kā labākais prognozēšanas modelis tiek atrasts modelis (1.9. izteiksme):

$$STARMA(S,p,q)x(P,Q)_D, \qquad (1.9.)$$

kur S – telpiskās atkarības kārta, $S=1$;

p – autoregresīvā polinoma kārta, $p=2$;

q – slīdošā vidējā polinoma kārta, $q=1$;

P – sezonālā autoregresīvā operatora kārta, $P=0$;

Q – sezonālā slīdošā vidējā operatora kārta, $Q=0$;

D – sezonālās komponentes periods, $D=24$.

Izmantojot modeli (1.9.), sasniegta vidējā kvadrātiskā kļūda (*Mean Square Error, MSE*) svārstās starp 1,962 un 2,101. Prognozēšanas precizitāte tiek salīdzināta ar *ARIMA* modeli, kurš dod *MSE* kļūdu starp 2,073 (par 5,7 % sliktāks rezultāts)

un 2,251 (par 7,1 % sliktāks rezultāts), līdz ar to *STARMA* modelis parāda labāku precizitāti, salīdzinot ar *ARIMA* modeli.

Pētījumā [61] *NDVI* indeksa laika rinda tiek prognozēta, izmantojot mākslīgos neironu tīklus. Pētījumu apgabals ietver četras dažādas ekoloģiskas zonas. Pētījumā izmantotā mākslīgā neironu tīkla modeļa piemērotie parametri tiek atrasti, izmantojot šķērsvalidāciju, un šī modeļa prognozēšanas precizitāte tiek salīdzināta ar Naivā Beijesa prognozēšanas modeļa precizitāti. Mākslīgo neironu tīkla modelis uzrāda augstu precizitāti, un tā precizitāte ir labāka par Naivā Beijesa modeļa precizitāti.

Lai veiktu pētījumu [40], tiek izmantoti *NOAA AVHRR NDVI* attēli, kas iegūti ar 14 dienu laika periodu. Papildus tiek izmantoti nokrišņu un temperatūras dati par to pašu laika periodu. Trīsdesmit viena skalārā *NDVI* laika rinda tiek iegūta, ņemot vidējo pikseļu vērtību *NDVI* attēlos, kur šie pikseļi bija 10 km diametrā ap meteoroloģiskajām stacijām. Kā prognozēšanas metode tiek izmantota automātiski regresējoša izplatītās aiztures (*Autoregressive Distributed-Lag, ADL*) metode, kur tiek izmantotas *NDVI* laika rindas vēsturiskās vērtības ar laika aizturi $\tau = 8$, tas ir, visas *NDVI* laika rindas vērtības no pēdējās zināmās vērtības (pašreizējā laika momentā t) līdz astotajai vērtībai pagātnē (laika momentā $t-7$), nokrišņu laika rindas vēsturiskās vērtības ar $\tau = 8$ un temperatūras laika rindas vēsturiskās vērtības ar $\tau = 6$. Prognoze tiek veikta vienu, divas, četras un sešas laika rindas vienības uz priekšu (attiecīgi divas, četras, astoņas un divpadsmit nedēļas uz priekšu). Iegūtā vidējā prognozēšanas precizitāte ir šāda: divu nedēļu prognozei determinācijas koeficients R^2 =0,966 un *MAPE* kļūda 5,0 %, četru nedēļu prognozei R^2 =0,931 un *MAPE* kļūda ir 7,4 %, astoņu nedēļu prognozei R^2 =0,854 un *MAPE* kļūda ir 11,1 % un divpadsmit nedēļu prognozei R^2 =0,798 un *MAPE* kļūda ir 12,9 %.

Vienvirziena neironu tīkla modelis tiek izmantots, lai prognozētu *NDVI* indeksa laika rindu pētījumā [60]. Pētījumu apgabals ietver Irānas ziemeļrietumus. Kā modeļa ieejas dati netiek izmantotas *NDVI* laika rindas vēsturiskās vērtības, bet tiek ņemtas divu klimatisku indeksu – Dienvidu svārstību indeksa (*Southern Oscillation Index*) un Ziemeļatlantijas svārstību indeksa (*North Atlantic Oscillation*) vēsturiskās vērtības. Tiek iegūta šāda prognozēšanas precizitāte: determinācijas koeficients R^2 =0,79, un saknes vidējā kvadrātiskā kļūda *RMSE*=0,011.

Pētījumā [37] *NDVI* indeksa laika rindu prognozēšanai tiek izmantota daudzparametru lineārā regresijas analīze. Izmantotie *NDVI* attēli ir nogludināti *NOAA*

AVHRR attēli. Kā papildu dati tiek izmantoti augsnes mitruma dati, zemes izmantojuma dati, kā arī jūras virsmas temperatūras dati. Pētījumu apgabals ir Indija. *NDVI* attēlu telpiskā izšķirtspēja ir 16 km, tādējādi tiek iegūtas daudzas *NDVI* laika rindas, atbilstoši pikseļu skaitam attēlos, kas noklāj visu Indijas teritoriju. Tiek pārbaudīti vairāki daudzparametru lineārās regresijas modeļi, kur modeļu ieejas datus veido:

- tikai *NDVI* indeksa vēsturiskās vērtības;
- augsnes mitruma un jūras virsmas temperatūras vēsturiskie dati;
- augsnes mitruma un jūras virsmas temperatūras vēsturiskie dati;
- augsnes mitruma, jūras virsmas temperatūras un *NDVI* indeksa vēsturiskie dati.

Prognoze tiek veikta vienu, divus un trīs mēnešus uz priekšu, kā labākais modelis tiek izvēlēts tas modelis, kuram determinācijas koeficienta R^2 vērtība starp prognozēto un novēroto *NDVI* indeksa laika rindu ir augstāka. Labāko precizitāti uzrādīja daudzparametru lineārās regresijas modelis ar trim ieejas parametriem (augsnes mitruma, jūras virsmas temperatūras un *NDVI* indeksa vēsturiskajiem datiem), un ar prognozēšanas periodu vienu mēnesi uz priekšu. Prognozētās *NDVI* indeksa vērtības arī tiek pārbaudītas uz zemes izmantojuma datiem, respektīvi, tiek pārbaudīts, vai prognozēšanas precizitāte ir pietiekami labā līmenī, arī atsevišķi katrai zemes izmantojuma kategorijai. Prognozēšanas precizitāte parāda, ka arī atsevišķi katrai zemes izmantojuma kategorijai, *NDVI* indeksa prognozēšanas precizitāte ir pietiekami laba.

Latvijā publiskā vidē pieejamie pētījumi, kas veltīti *NDVI* indeksa laika rindu prognozēšanai, nav izplatīti. Šāds *NDVI* indeksa prognozēšanas pētījumu trūkums Latvijas reģionālajā līmenī nozīmē, ka ir pieejama brīva pētniecības sadaļa, kurā ir iespējams veikt pētījumus, attīstot šo pētniecības sadaļu Latvijā un reizē attīstot arī precīzās lauksaimniecības nozari Latvijā.

NDVI laika rindu prognozēšanas pētījumos [5], [13], [24], [38], [40], [60], [61] izplatītāko prognozēšanas metožu: *ARIMA* modeļa, daudzparametru lineārās regresijas analīzes un vienvirziena neironu tīklu salīdzinājums pēc to piemērotības nelineāru, nestacionāru un trokšņainu laika rindu prognozēšanai dots 1.4. tabulā. Atbilstošā laukā, kur krustojas konkrētā prognozēšanas metode un atbilstošā īpašība, tiek pielikts “+”, ja šai metodei piemīt šī īpašība, bet, ja nepiemīt, tad tiek pielikts “– ”. Tā kā *ARIMA* metode un daudzparametru lineāra regresijas analīze ir lineārās prognozēšanas

metodes, tad tās pēc savas būtības nav piemērotas nelineāru laika rindu prognozēšanai, un nav robustas pret trokšņiem laika rindās [112]. Tomēr *ARIMA* metodi iespējams pielietot nestacionāru laika rindu prognozēšanai, pateicoties laika rindu diferencēšanai [3].

1.4. tabula

NDVI laika rindu prognozēšanā izmantoto metožu salīdzinājums

Metodes & algoritmi / **Īpašības**	***ARIMA* metode**	**Daudzparametru lineārā regresijas analīze**	**Vienvirziena neironu tīkli**
Piemērotība nelineāru procesu prognozēšanai	–	–	+
Piemērotība nestacionāru procesu prognozēšanai	+	–	+
Robustums pret trokšņiem	–	–	+

Augstāk esošajā tabulā (skat. 1.4. tab.) var redzēt, ka literatūras apskatā analizētajos pētījumos starp populārākajām prognozēšanas metodēm vispiemērotākā prognozēšanas metode *NDVI* laika rindām ir vienvirziena neironu tīkli, kas ir nelineāra prognozēšanas metode un labi spēj prognozēt nelineāras [62], [114], nestacionāras [14], [47] un trokšņainas [15], [20] laika rindas.

1.9.2. NDVI laika rindu prognozēšanas pētījumu kopsavilkums

Kā redzams esošajos *NDVI* indeksa laika rindu prognozēšanas pētījumu rezultātos, pārsvarā tiek izmantotas Boksa–Dženkinsa *ARIMA* metode, mākslīgie neironu tīkli un lineārā regresijas analīze. Vairāki prognozēšanas modeļi kā ieejas datus izmanto ne tikai pašas *NDVI* indeksa laika rindas vēsturiskās vērtības, bet arī papildu datus [38], [40], [60]. Analizētajos zinātniskajos rakstos iegūta dažāda prognozēšanas precizitāte, kas uzskatāma gan kā laba, un ļauj izmantot šos prognozēšanas modeļus praktisku uzdevumu risināšanā, piemēram, pētījumos [24], [40] iegūtais determinācijas koeficients $R^2 = 0,96$, gan arī kā neapmierinoša precizitāte, kura nav pietiekama praktisku problēmu risināšanai, proti, pētījumā [5] iegūtais korelācijas koeficients $R = 0,077$.

Precīzās *NDVI* indeksa laika rindu prognozes var palīdzēt veikt nepieciešamos aprēķinus precīzajā lauksaimniecībā, bet no prognožu kļūdu lieluma ir atkarīgs, piemēram, cik daudz būs lieki vai nepietiekoši iztērētu barības vielu, un attiecīgi, kādi būs zaudējumi. Tātad, lai to izmantotu precīzajā lauksaimniecībā, ir nepieciešama prognozēšanas sistēma, kas spēj sasniegt augstu *NDVI* indeksa laika rindu prognozēšanas precizitāti. Apskatītajos pētījumos ne vienmēr tiek sasniegta augsta prognozēšanas precizitāte, tā ir atkarīga no dažādiem faktoriem, piemēram, no papildu datu izmantošanas.

Vairākos pētījumos tika izmantoti papildu ieejas dati, piemēram, gaisa temperatūras vai nokrišņu dati. Papildu datu iegūšana var būt sarežģīta, jo:

- ne vienmēr konkrētajam pikselim ir pieejami šie dati, biežāk šie dati ir pieejami par lieliem apgabaliem (vidējās vērtības novadā, rajonā utt.);
- ja *NDVI* attēlus var iegūt bez maksas, tad šie papildu dati var būt maksas dati.

Praktiski visos apskatītajos pētījumos maz uzmanības tiek pievērsts datu pirmapstrādes posmam, netiek izmantotas tādas pirmapstrādes metožu grupas kā pazīmju atlasīšana (*feature selection*) un pazīmju izvilkšana (*feature extraction*), kas palīdz sagatavot laika rindu prognozēšanas uzdevumos ieejas datu kopu tā, lai tiktu paaugstināta prognozēšanas precizitāte.

Boksa–Dženkinsa *ARIMA* metode, kā arī lineārā regresijas analīze, kas tiek izmantota *NDVI* laika rindu prognozēšanai vairākos pētījumos, ir lineārās prognozēšanas metodes, un nav robustas pret trokšņiem, savukārt *NDVI* laika rindas ir nelineāras un trokšņainas. Turklāt piemērota *ARIMA* modeļa noteikšana ir laikietilpīga un resursus patērējoša procedūra.

Pēdējos gados laika rindu prognozēšanas uzdevumos tiek plaši izmantotas signālu dekompozīcijas metodes, kas balstītas uz spektrālo analīzi. Šīs metodes laika rindu prognozēšanas uzdevumos tiek izmantotas, lai no vienas laika rindas iegūtu vairākas atvasinātas laika rindas un, padodot prognozēšanas metodei ieejā ne tikai *NDVI* laika rindas vērtības, bet arī atvasināto laika rindu vērtības, sasniegtu augstāku prognozēšanas precizitāti nekā bez šo metožu izmantošanas [30], [33], [95]. Lai varētu analizēt, vai signālu dekompozīcijas metodes būtu pielietojamas *NDVI* laika rindu prognozēšanas uzdevumā, ir jāveic laika rindu prognozēšanas pētījumu analīzi, kuros tiek pielietotas uz spektrālo analīzi balstītas dekompozīcijas metodes. Tas tiek izdarīts 1.10. apakšnodaļā.

1.10. Laika rindu prognozēšanas pētījumu analīze ar dekompozīcijas metožu pielietojumu

Laika rindu dekompozīcija ļauj sadalīt laika rindu vairākās komponentēs, iegūstot atvasinātas laika rindas. Tradicionāli plaši izmantota ir laika rindu dekompozīcija trendā, cikliskajā komponentē, sezonālajā komponentē un neregulārajā komponentē (troksnī), izmantojot aditīvo vai multiplikatīvo modeli. Katra no šīm komponentēm tiek prognozēta atsevišķi, un tad prognozes tiek summētas vai reizinātas, atkarībā no izmantotā modeļa, un tiek iegūta oriģinālās laika rindas prognoze. Bet laika rindu dekompozīcija tiek veikta, izmantojot arī digitālu signālu apstrādē izmantotās dekompozīcijas metodes, kas balstītas uz spektrālo (frekvenču) analīzi [30], [33], [34], [63], [84], [85], [95].

1.10.1. Prognozēšanas pētījumu analīze ar dekompozīcijas lietošanu

Veicot zinātniskās literatūras analīzi, tiek secināts, ka izplatītākās uz spektrālo analīzi balstītās dekompozīcijas metodes, ko izmanto laika rindu prognozēšanas uzdevumos, ir variāciju apakšsignālu dekompozīcijas (*Variational Mode Decomposition, VMD*) metode (pētījumu apskats dots 1.5. tabulā), veivletu dekompozīcijas (*Wavelet Decomposition, WD*) un empīrisko apakšsignālu dekompozīcijas (*Empirical Mode Decomposition, EMD*) metodes (pētījumu apskats dots 1.6. tabulā). Šīs metodes tiek izmantotas datu pirmapstrādes posmā, sadalot laika rindu vairākās komponentēs jeb apakšsignālos, kur šo apakšsignālu summa veido oriģinālo laika rindu.

Pētījumā [85] tiek veikta divu jēlnaftas cenu: *Brent Crude Oil Spot* (*Brent*) un *West Texas Intermediate* (*WTI*) prognozēšana, izmantojot veivletu dekompozīcijas un vienvirziena neironu tīkla (*WD-BPFNN*) modeli. Vispirms ieejas laika rinda, izmantojot veivletu dekompozīciju, tiek sadalīta vairākās komponentēs. Tad noteikts skaits šo komponenšu pagātnes vērtību, kas atbilst loga garumam, tiek padotas vienvirziena neironu tīkla ieejā, un tiek prognozēta jēlnaftas cenas vērtība vienu laika soli uz priekšu. Prognozēšana arī tiek veikta, izmantojot tikai vienvirziena neironu tīklu.

Urāna resursu cenas prognozēšana tiek veikta pētījumā [110]. Oriģinālā laika rinda, izmantojot *EMD* metodi, tika sadalīta vairākās komponentēs – laika rindās. Katrai no šīm laika rindām tiek veikta fāžu telpas rekonstrukcija (*Phase Space Reconstruction, PSR*) izmantojot fāžu telpas rekonstrukcijas metodi ar laika aizturi.

1.5. tabula

Laika rindu prognozēšanas pētījumu apskats ar *VMD* metodes izmantošanu

<table>
<tr><th>Dekompo-zīcijas metode</th><th>Prognozē-šanas metode</th><th colspan="4">Kļūdu un precizitātes rādītāji</th></tr>
<tr><td rowspan="2">*VMD* [30]</td><td rowspan="2">*BPFNN*; *ARMA*</td><td colspan="2">VMD-BPFNN modelim:</td><td>BPFNN modelim:</td><td>ARMA modelim</td></tr>
<tr><td colspan="2">*RMSE*=0,07.</td><td>*RMSE*=0,27.</td><td>*RMSE*=0,34.</td></tr>
<tr><td rowspan="2">*VMD*; *EMD* [95]</td><td rowspan="2">*SNN*; *BPFNN*</td><td>VMD-SNN modelim:</td><td>VMD-BPFNN modelim:</td><td>EMD-SNN modelim:</td><td>SNN modelim:</td></tr>
<tr><td>*RMSE*=0,04; *MAE*=0,03; *MAPE*=2,19 %; *R*=0,99.</td><td>*RMSE*=0,12; *MAE*=0,09; *MAPE*=6,49 %; *R*=0,98.</td><td>*RMSE*=0,15; *MAE*=0,12; *MAPE*=10,54 %; *R*=0,97.</td><td>*RMSE*=0,20; *MAE*=0,16; *MAPE*=10,82 %; *R*=0,96.</td></tr>
<tr><td rowspan="2">*VMD*, *EMD* [34]</td><td rowspan="2">*WRELM*</td><td colspan="2">EMD-PACF-WRELM</td><td colspan="2">VMD-PACF-WRELM</td></tr>
<tr><td>*MAPE*=16,63 %.</td><td>*RMSE*=0,55.</td><td>*MAPE*=11,09 %.</td><td>*RMSE*=0,25.</td></tr>
<tr><td rowspan="4">*VMD*, *WD* [84]</td><td rowspan="4">*ANN*, *ELM*, *LSSVM*</td><td colspan="2">VMD-ANN</td><td colspan="2">VMD-ELM</td><td>VMD-LSSVM</td></tr>
<tr><td colspan="2">*RMSE*=6,06; *MAE*=3,00; R^2=0,99.</td><td colspan="2">*RMSE*=3,19; *MAE*=1,37; R^2=0,99.</td><td>*RMSE*=2,88; *MAE*=1,41; R^2=0,99.</td></tr>
<tr><td colspan="2">WD-ANN</td><td colspan="2">WD-ELM</td><td>WD-LSSVM</td></tr>
<tr><td colspan="2">*RMSE*=7,09; *MAE*=3,47; R^2=0,98.</td><td colspan="2">*RMSE*=4,20; *MAE*=1,69; R^2=0,99.</td><td>*RMSE*=4,40; *MAE*=1,66; R^2=0,99.</td></tr>
</table>

Katrs fāžu telpu datu paraugs tiek normalizēts intervālā [0; 1]. Tālāk, izmantojot normalizētās fāžu telpas kā ieejas datus, tiek prognozēts katrs no apakšsignāliem vienu laika soli uz priekšu. Apakšsignālu prognozēšanai tiek izmantota ekstrēma apmācības mašīna (*Extreme Learning Machine*, *ELM*) un beigās visas prognozes tiek summētas, iegūstot urāna resursu cenas prognozi. Šis prognozēšanas modelis apzīmēts kā *EMD-PSR-ELM* modelis. Tiek arī veikta prognozēšana, izmantojot *PSR-ELM* modeli un *PSR-RBFNN* modeli, kur ar *RBFNN* apzīmēts radiālās bāzes funkciju neironu tīkls (*Radial Basis Function Neural Network*, *RBFNN*).

1.6. tabula

Laika rindu prognozēšanas pētījumu apskats ar *WD* un *EMD* metožu izmantošanu

<table>
<tr><th>Dekom-pozīcijas metode</th><th>Prognozē-šanas metode</th><th colspan="8">Kļūdu un precizitātes rādītāji</th></tr>
<tr><td rowspan="2">*WD* [85]</td><td rowspan="2">*BPFNN*</td><td colspan="6">WD-BPFNN modelim:</td><td colspan="2">BPFNN modelim:</td></tr>
<tr><td colspan="6">*Brent* cenai *MAPE*=1,31 %; *RMSE*=0,18;
WTI cenai *MAPE*=1,39 %; *RMSE*=0,74.</td><td colspan="2">*Brent* cenai *MAPE*=1,64 %; *RMSE*=0,99;
WTI cenai *MAPE* 1,79 % un *RMSE* 0,94.</td></tr>
<tr><td rowspan="2">*EMD* [110]</td><td rowspan="2">*ELM*; *RBFNN*</td><td colspan="2">EMD-ELM modelim</td><td colspan="4">ELM modelim</td><td colspan="2">RBFNN modelim</td></tr>
<tr><td colspan="2">*RMSE*=0,82; *MAE*=0,69;
DS=80 %.</td><td colspan="4">*RMSE*=2,23; *MAE*=1,47; *DS*=68 %.</td><td colspan="2">*RMSE*=3,44; *MAE*=2,42; *DS*=44 %.</td></tr>
<tr><td rowspan="2">*EMD*; *WD* [33]</td><td rowspan="2">*BPFNN*</td><td colspan="2">EMD-PSO-BPFNN modelim:</td><td colspan="2">WD-PSO-BPFNN modelim:</td><td colspan="2">EMD-GA-BPFNN modelim:</td><td colspan="2">PSO-BPFNN modelim:</td><td colspan="2">**GA-*BPFNN* modelim:**</td></tr>
<tr><td colspan="2">*R*=0,99;
RMSE=3,16e+3;
MAPE=0,02 %.</td><td colspan="2">*R*=0,97;
RMSE=8,15e+3;
MAPE=0,13 %.</td><td colspan="2">*R*=0,94;
RMSE=12,66e+3;
MAPE=0,34 %.</td><td colspan="2">*R*=0,99;
RMSE=3,77e+3;
MAPE=0,11 %.</td><td colspan="2">*R*=-0,06;
RMSE=36,82e+3;
MAPE=1,13 %.</td></tr>
<tr><td rowspan="3">*EMD* [63]</td><td rowspan="3">*SARIMA*; Holta-Vintersa modelis; Naivais Beijesa modelis</td><td colspan="4">Prognozējot laika rindu ar sešu mēnešu periodu</td><td colspan="4">Prognozējot laika rindu ar divpadsmit mēnešu periodu</td></tr>
<tr><td>EMD-SARIMA modelim:</td><td>SARIMA modelim:</td><td>Holta-Vintersa modelim:</td><td>Naivajam Beijesa modelim:</td><td>EMD-SARIMA modelim</td><td>SARIMA modelim:</td><td>Holta-Vintersa modelim:</td><td>Naivajam Beijesa modelim:</td></tr>
<tr><td>*MAE*=5,02;
MAPE=1,42 %.</td><td>*MAE*=15,80;
MAPE=4,62 %.</td><td>*MAE*=14,30;
MAPE=4,20 %.</td><td>*MAE*=21,33
MAPE=6,34 %.</td><td>*MAE*=15,94;
MAPE=4,64 %.</td><td>*MAE*=31,90;
MAPE=9,94 %.</td><td>*MAE*=26,02;
MAPE=8,33 %.</td><td>*MAE*=44,02;
MAPE=13,5 %.</td></tr>
</table>

Empīriskā apakšsignālu dekompozīcijas metode un vienvirziena neironu tīkls (skat. 1.6. tabulu), optimizēts ar daļiņu spieta optimizācijas *(Particle Swarm Optimization, PSO*) metodi, tiek izmantoti ambulatoro apmeklējumu laika rindas prognozēšanai [33]. Oriģinālā laika rinda tiek sadalīta vairākās komponentēs, kur katra komponente tiek prognozēta. Prognozēšana tiek veikta, izmantojot vienvirziena neironu tīklu, un prognozes tiek summētas, iegūstot ambulatoro apmeklējumu prognozi. Papildu prognozēšana tiek veikta arī ar:

- *WD-PSO-BPFNN* modeli (veivletu dekompozīcija (*WD*) un kļūdas atgriezeniskās izplatības vienvirziena neironu tīkls (*Back Propagation Feed Forward Neural Network, BPFNN*) optimizēts ar daļiņu spieta optimizācijas metodi (*PSO*);
- *EMD-GA-BPFNN* modeli (empīriskā apakšsignālu dekompozīcijas metode (*EMD*) un kļūdas atgriezeniskās izplatības vienvirziena neironu tīkls (*BPFNN*), optimizēts ar ģenētisko algoritmu (*Genetic Algorithm, GA*));
- *PSO-BPFNN* modeli (kļūdas atgriezeniskās izplatības vienvirziena neironu tīkls (*BPFNN*), optimizēts ar daļiņu spieta optimizācijas metodi (*PSO*));
- *GA-BPFNN* (kļūdas atgriezeniskās izplatības vienvirziena neironu tīkls (*BPFNN*), optimizēts ar ģenētisko algoritmu (*GA*)).

Lai veiktu gaisa satiksmes kravu trafika laika rindu ar sešu mēnešu un divpadsmit mēnešu periodu prognozēšanu pētījumā [63], empīriskā apakšsignālu dekompozīcijas metode tiek kombinēta ar sezonālo *ARIMA* (*SARIMA*) modeli, iegūstot *EMD-SARIMA* modeli. Līdzīgi kā vairākos iepriekš apskatītos pētījumos, oriģinālā gaisa satiksmes trafika laika rinda tiek sadalīta vairākās komponentēs, katra komponente tika prognozēta, izmantojot *SARIMA*, un beigās prognozes tiek summētas. Prognozēšana arī tiek veikta, izmantojot *SARIMA* modeli bez dekompozīcijas veikšanas, Holta-Vintersa modeli un Naivo Beijesa modeli bez dekompozīcijas izmantošanas.

Mājsaimniecību elektroenerģijas patēriņa laika rinda pētījumā [30] tiek prognozēta, izmantojot *VMD-BPFNN* modeli. Vispirms oriģinālajai laika rindai tiek veikta fāžu telpas rekonstrukcija ar laika aiztures metodi, tad katram aiztures vektoram no rekonstruētās fāžu telpas, kas atbilst oriģinālās laika rindas fragmentam ar noteiktu loga garumu, tiek iegūts noteikts skaits apakšsignālu, izmantojot variāciju apakšsignālu dekompozīcijas metodi. Visi iegūtie apakšsignāli tiek padoti ieejā kļūdas atgriezeniskās

izplatības vienvirziena neironu tīklam, un tiek veikta elektroenerģijas patēriņa laika rindas prognoze vienu laika soli uz priekšu. Prognozēšana arī tiek veikta, izmantojot vienvirziena neironu tīklu bez dekompozīcijas veikšanas un autoregresīvo slīdošā vidējā (*Autoregressive Moving Average, ARMA*) modeli.

Variāciju apakšsignālu dekompozīcijas metode (*VMD*) un impulsu neironu tīkls (*Spiking Neural Network, SNN*) tiek izmantoti oglekļa cenas laika rindas prognozēšanai pētījumā [95]. Oriģinālā laika rinda, izmantojot variāciju apakšsignālu dekompozīcijas metodi, tiek sadalīta vairākās komponentēs. Katra komponente tiek prognozēta, izmantojot virsotņu neironu tīklu, un prognozes tiek summētas, iegūstot oglekļa cenas prognozi (*VMD-SNN* modelis). Prognozēšana arī tiek veikta, izmantojot, *VMD-BPFNN* modeli, *EMD-SNN* modeli, virsotņu neironu tīklu bez dekompozīcijas veikšanas un kļūdas atgriezeniskās izplatības vienvirziena neironu tīklu bez dekompozīcijas veikšanas.

Latvijā ir tapuši vairāki promocijas darbi, kuros tiek apskatīta un analizēta signālu dekompozīcija [97], [102], kas balstīta uz spektrālo analīzi. Tomēr Latvijā veiktie pētījumi par laika rindu prognozēšanu, izmantojot uz spektrālo analīzi balstītās signālu dekompozīcijas metodes, publiskā vidē nav izplatīti.

Vairāku pētījumu analīze dota 1.7. tabulā, kur tiek izmantota kāda no trim (*VMD, EMD, WD*) uz spektrālo analīzi balstītām signālu dekompozīcijas metodēm. Zemāk esošajā 1.7. tabulā kolonnas satur trīs dekompozīcijas metodes: veivletu dekompozīcijas metodi [33], [84], empīrisko apakšsignālu dekompozīcijas metodi [33], [34], [51], [64], [95] un variāciju apakšsignālu dekompozīcijas metodi [34], [51], [64], [84], [95], savukārt rindas satur attiecīgās prognozēšanas metodes, kas tiek izmantotas dažādos pētījumos. Par prognozēšanas metodēm šajos sešos pētījumos tiek izmantoti: vienvirziena neironu tīkls [33], mazāko kvadrātu atbalsta vektoru mašīna [84], ekstrēmās apmācības mašīna [84], svērta regularizētā ekstrēma apmācības mašīna [34], impulsu neironu tīkls [95], vispārējās regresijas neironu tīkls [51] un daudzkodolu regularizēts pseido-inversais neironu tīkls [64]. Turklāt katrā pētījumā tiek izmantotas divas no šīm laika rindu dekompozīcijas metodēm, tādā veidā nodrošinot prognozēšanas precizitātes salīdzināšanas iespēju: vispirms prognozēšanas metode tiek izmantota kopā ar pirmo dekompozīcijas metodi, tad – ar otro dekompozīcijas metodi. Tai dekompozīcijas metodei, ko izmantojot tiek sasniegta augstākā prognozēšanas precizitāte, atbilstoša laukā, kur krustojas šī dekompozīcijas metode un atbilstošā prognozēšanas metode, tiek pielikts "+", otrai dekompozīcijas

metodei, ko izmanto pētījumā, atbilstošajā laukā tiek pielikts "–", kā redzams 1.7. tabulā. Trešā dekompozīcijas metode šajos pētījumos netiek izmantota.

1.7. tabula

Laika rindu prognozēšanā izmantoto dekompozīcijas metožu salīdzinājums

Dekompozīcijas metode / Prognozēšanas metode	Veivletu dekompozīcija	Empīriskā apakšsignālu dekompozīcija	Variāciju apakšsignālu dekompozīcija
Vienvirziena neironu tīkls [33]	–	+	
Mazāko kvadrātu atbalsta vektoru mašīna, ekstrēmā apmācības mašīna [84]	–		+
Svērta regularizētā ekstrēma apmācības mašīna [34]		–	+
Impulsu neironu tīkls [95]		–	+
Vispārējs regresijas neironu tīkls [51]		–	+
Daudzkodolu regularizēts pseido–inversais neironu tīkls [64]		–	+

Tādēļ lauks, kur krustojas trešā dekompozīcijas metode un atbilstošā prognozēšanas metode, tiek aizkrāsots ar pelēku krāsu. Ir redzams, ka visbiežāk augstākā laika rindu prognozēšanas precizitāte tiek sasniegta, izmantojot *VMD* metodi. *VMD* metode četras reizes ir palīdzējusi sasniegt augstāku prognozēšanas precizitāti, salīdzinot ar *EMD* metodi, un vienreiz, salīdzinot ar *WD* metodi. Savukārt vienā pētījumā *EMD* metode ir palīdzējusi sasniegt augstāku prognozēšanas precizitāti, salīdzinot ar *WD* metodi.

1.10.2. Dekompozīcijas metožu izmantošanas pētījumu kopsavilkums

Kā redzams 1.10. apakšnodaļā apskatītajos pētījumos, dekompozīciju metodes izmantošana laika rindu prognozēšanas pētījumos noved pie mazāku prognozēšanas kļūdu sasniegšanas, nekā neizmantojot laika rindu dekompozīciju. Tāpat šajos

pētījumos var redzēt, ka variāciju apakšsignālu dekompozīcijas metodes izmantošana iepriekš apskatītajos pētījumos ļauj iegūt mazākas prognozēšanas kļūdas, salīdzinot ar citām dekompozīcijas metodēm, kas balstītas uz spektrālo analīzi. No tā var secināt, ka apskatītajos pētījumos, variāciju apakšsignālu dekompozīcijas metode spēj dot visaugstāko pienesumu prognozēšanas precizitātes paaugstināšanā, salīdzinot ar citām dekompozīcijas metodēm, un tādēļ būtu arī izmantojama darbā izstrādājamajā sistēmā.

Apskatītajos pētījumos 1.10. apakšnodaļā daži autori [33], [95] lieto dekompozīcijas metodi visai ieejas laika rindai, un tikai tad tiek izveidotas datu kopas, un tiek veikta prognozēšanas modeļa apmācība. Iegūtie rezultāti parāda augstu prognozēšanas precizitāti, tomēr šeit tiek prognozēti tikai vēsturiskie novērojumi, kur testa datu kopu veido daļa no šiem vēsturiskajiem novērojumiem, uz kuriem pirms tam tika pielietota dekompozīcija. Bet reālajā laikā nāk klāt jaunas laika rindas vērtības, uz kurām netika izmantota dekompozīcijas metode, un to prognozēšana šajos zinātniskajos rakstos [33], [95] nav apskatīta. Līdz ar to nav iespējams noteikt, vai šie apmācītie prognozēšanas modeļi spēj ar līdzvērtīgu precizitāti prognozēt arī jaunas laika rindas vērtības, kuras nāks klāt laika rindai aiz pēdējā vēsturisko novērojumu elementa.

Savukārt citi autori sadala ieejas laika rindu mazākos fragmentos, atbilstoši kādam loga garumam un lieto dekompozīciju katram fragmentam atsevišķi [30], [34]. Praktiski tas nozīmē, ka dekompozīcija tiek lietota katram datu ierakstam no rekonstruētās fāžu telpas. Šāda veida uz spektrālo analīzi balstītas signālu dekompozīcijas metodes var izmantot arī jaunu laika rindas vērtību prognozēšanai (to vērtību prognozēšanai, kas nāk pēc vēsturiskiem datiem).

Tomēr visu pētījumos apskatīto signālu dekompozīcijas metožu galvenais trūkums ir tas, ka tās nav paredzētas laika rindu dekompozīcijai prognozēšanas uzdevumu ietvaros. Pirmā problēma ir, ka, aprēķinot apakšsignālu vērtības kādā laika solī, dekompozīcijas metodes izmanto visas ieejas laika rindas vērtības, tajā skaitā arī nākotnes novērojumu vērtības, skatoties pēc laika. Tādējādi, ja dekompozīcijas metode izmanto visus ieejas laika rindas vēsturiskos novērojumus, tad katru reizi, nākot klāt jaunam novērojumam, apakšsignālu vērtības ir jāpārrēķina vēlreiz, dekompozīcijas metodei padodot ieejā vēsturiskos novērojumus kopā ar jaunāko novērojumu, un visas iepriekš aprēķinātās apakšsignālu vērtības visos laika soļos mainās. Otrā problēma ir tā, ka, lietojot dekompozīcijas metodi visai ieejas laika rindai, netiek pārbaudīts, vai ir iespējams prognozēt jaunu laika rindas vērtību ar līdzvērtīgu precizitāti, salīdzinot ar to, kas iegūta, prognozējot vēsturiskos *NDVI* novērojumus.

1.11. Prognozēšanas sistēmas konceptuālais apraksts

Pamatojoties uz uzdevuma formālo nostādni, analizējamo datu struktūru un abiem literatūras apskatiem, tiek piedāvāta prognozēšanas sistēmas konceptuālā shēma, kas ietver datu pirmapstrādes, signālu apstrādes un mašīnmācīšanās metodes un pieejas (skat. 1.5. att.). Balstoties uz 1.9. un 1.10. apakšnodaļās veikto literatūras apskatu un analīzi, ir izvirzītas prognozēšanas sistēmas **prasības**:

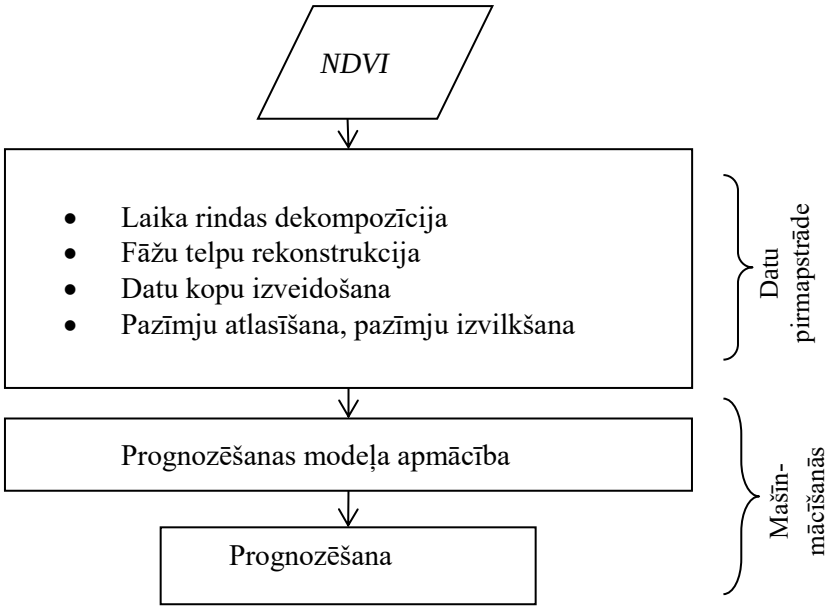
1. prognozēšanas sistēmai ieejā jāsaņem tikai *NDVI* laika rindas vēsturiskie dati;
2. jābūt izstrādātai pieejai, kas ļauj aproksimēt apakšsignālu vērtības katrā laika solī, kas nodrošina to, ka:
   - apakšsignālu vērtības ir pieejamas visiem laika soļiem, kur ir pieejami vēsturiskie normalizēta veģetācijas indeksa laika rindu novērojumi;
   - katrā laika solī apakšsignālu vērtības tiek iegūtas bez variāciju apakšsignālu dekompozīcijas metodes izmantošanas;
   - aprēķinot apakšsignālu vērtības katrā jaunā laika solī, iepriekšējos laika soļos aprēķinātās apakšsignālu vērtības nemainās;
3. prognozēšanas precizitātes uzlabošanai prognozēšanas modelim ir jābūt kombinētam:
   - ar datu pirmapstrādes metodēm (pazīmju atlasīšana un izvilkšana);
   - ar piedāvāto apakšsignālu aproksimācijas pieeju.

Izstrādājot prognozēšanas sistēmu un balstoties uz izvirzītajām prasībām, pastāv vairāki pētījuma **ierobežojumi**.

1. Tiek analizētas un apstrādātas *NDVI* laika rindas, kur katra iegūta no viena pikseļa.
2. Tiek veikta tikai īstermiņa prognozēšana – vienu laika rindas vienību uz priekšu (promocijas darbā izmantotajiem *MODIS Terra NDVI* attēliem tās ir septiņas dienas).
3. Netiek izmantoti papildu dati (piemēram, gaisa temperatūras dati, nokrišņu dati, citi veģetācijas indeksi, zemes virsmas kategoriju dati).

Analizējamajiem datiem, kas sastāv no ieejas laika rindas, tiek piedāvāts veikt datu pirmapstrādi. Datu pirmapstrādes procesā, izmantojot dekompozīcijas metodi, tiek piedāvāts iegūt noteiktu skaitu laika rindu. Gan ieejas laika rindai, gan atvasinātajām

laika rindām tiek piedāvāts rekonstruēt fāžu telpas, sadalot šīs laika rindas fragmentos pēc kādas noteiktas fāžu telpas dimensijas (loga garuma). Fāžu telpas tiek apvienotas, iegūstot sākotnējo ieejas datu kopu, kas var tikt sadalīta apmācības, validācijas un testa datu kopās, kur katra no šīm trim kopām vēl tiek sadalīta atribūtu jeb pazīmju kopā un prognozējamā parametrā.



1.5. att. Prognozēšanas sistēmas konceptuālā shēma.

Tiek piedāvāts izmantot pazīmju atlasīšanu, lai noteiktu atribūtu informatīvumu un izslēgtu no atribūtu kopām neinformatīvos atribūtus, kas var mazināt prognozēšanas precizitāti. Reducētām atribūtu kopām tiek piedāvāts izmantot pazīmju izvilkšanu, transformējot kopas pēc noteiktiem kritērijiem. Tālāk apmācības un validācijas datu kopas tiek padotas ieejā prognozēšanas metodei, kas, apmācoties uz šīm datu kopām, atrod funkcionālo sakarību (1.7, iegūstot prognozēšanas modeli, ko var izmantot jauno ieejas laika rindas vērtību prognozēšanai.

1.12. Kopsavilkums un secinājumi

Nodaļas ietvaros ir sniegts ieskats *NDVI* indeksa teorētiskajos pamatos, šī indeksa laika rindu prognozēšanas pētījumos un dažādu laika rindu prognozēšanas pētījumos, kur tiek izmantotas uz spektrālo analīzi balstītas signālu dekompozīcijas metodes. Definēta problēmvide un prognozēšanas uzdevuma formālā nostādne. Sniegts izstrādājamās prognozēšanas sistēmas konceptuālais apraksts. Izdarīti galvenie secinājumi par veikto literatūras apskatu un analīzi.

1. Apzinātas galvenās problēmas, kas saistītas ar *NDVI* indeksa laika rindu prognozēšanu: zema prognozēšanas precizitāte, nepietiekams datu pirmapstrādes metožu pielietojums, lineāru prognozēšanas modeļu izmantošana un ārējo datu (bieži vien maksas datu) izmantošana.
2. Apzinātas galvenās problēmas uz spektrālo analīzi balstīto dekompozīcijas metožu izmantošanai laika rindu prognozēšanas uzdevumos. Pirmā problēma ir, ka, aprēķinot apakšsignālu vērtības kādā laika solī, dekompozīcijas metodes izmanto visas ieejas laika rindas vērtības, tajā skaitā arī nākotnes novērojumu vērtības, skatoties pēc laika. Tādējādi, ja dekompozīcijas metodei ieejā tiek padoti visi laika rindas vēsturiskie novērojumi, tad, nākot klāt jaunam novērojumam, katru reizi apakšsignālu vērtības ir jāpārrēķina pa jaunu, un visas iepriekš aprēķinātās apakšsignālu vērtības visos laika soļos mainās.
3. Otrā problēma ir, ka, pielietojot dekompozīcijas metodi visai ieejas laika rindai, vairākos pētījumos, kur izmanto šādu pieeju, netiek pārbaudīts, vai ir iespējams prognozēt jaunu laika rindas vērtību ar līdzvērtīgu precizitāti, salīdzinot ar to, kas iegūta, prognozējot vēsturiskos *NDVI* laika rindas novērojumus.

Nodaļā arī ir iezīmētas metožu grupas, kas būtu pielietojamas *NDVI* indeksa laika rindu pirmapstrādei un prognozēšanai, izstrādājot *NDVI* indeksa laika rindu prognozēšanas sistēmu. Tiek piedāvāta shēma ar laika rindas dekompozīciju, fāžu telpu rekonstrukciju visām laika rindām, fāžu telpu apvienošanu, pazīmju atlasīšanu, pazīmju izvilkšanu un prognozēšanas metodes izmantošanu.

2. PROGNOZĒŠANAS SISTĒMAS IZSTRĀDĀŠANĀ IZMANTOJAMO METOŽU APSKATS

Viena no mūsdienu tendencēm, izstrādājot prognozēšanas sistēmas, ir dažādu metožu un pieeju kombinācija, parasti kombinējot dažādas datu pirmapstrādes metodes ar vienu vai vairākām laika rindu prognozēšanas metodēm [35], [80], [81], [85], [110]. Tas ļauj kompensēt atsevišķu metožu vai pieeju trūkumus un paaugstināt prognozēšanas precizitāti. Šī nodaļa ir veltīta populāru datu pirmapstrādes un laika rindu prognozēšanas metožu apskatam, kas tiek izmantotas izstrādājamajā prognozēšanas sistēmā.

Atbilstoši literatūras analīzei 1.10. apakšnodaļā, par laika rindu dekompozīcijas metodi tiek piedāvāts izmantot variāciju apakšsignālu dekompozīcijas (*Variational Mode Decomposition, VMD*) metodi [118]. Lai skalāru normalizētā veģetācijas indeksa (*Normalized Difference Vegetation Index, NDVI*) laika rindu sadalītu fragmentos pēc noteikta loga garuma un iegūtu sākotnējo datu kopu, var tikt izmantota fāžu telpas rekonstrukcija ar laika aiztures metodi [49]. Sākotnējā datu kopa tiek sadalīta apmācības, validācijas un testa datu kopās, kur katra no šīm kopām sastāv no atribūtu jeb pazīmju kopas un prognozējamā parametra. Katrs atribūts šajās kopās ir pa noteiktu laika soli nobīdītas *NDVI* laika rindas un apakšsignālu vērtības. Lai reducētu sākotnējās atribūtu kopas, atlasot tikai informatīvākos atribūtus, var izmantot pazīmju atlasīšanas metodes [26], [77]. Populāra pazīmju atlasīšanas metode, ko var lietot laika rindu prognozēšanas uzdevumos, ir pakāpeniskā regresijas analīze [98], ko tiek piedāvāts izmantot izstrādājamā sistēmā. Ir svarīgi, lai atribūti savstarpēji lineāri nokorelētu un neapgrūtinātu prognozēšanas modeļa apmācību. Principiālo komponentu analīze (*Principal Component Analysis, PCA*) ir metode, ko bieži izmanto lineāri nekorelējošu atribūtu kopu iegūšanai [79]. Principiālo komponentu analīzes metodi tiek piedāvāts izmantot prognozēšanas sistēmā kā pazīmju izvilkšanas metodi. No prognozēšanas metodēm sistēmā tiek piedāvāts izmantot rekurento slāņu neironu tīklus (*Layer Recurrent Neural Network, LRNN*) pateicoties savām īpašībām labi prognozēt nelineāras, nestacionāras un trokšņainas laika rindas [71].

2.1. Variāciju apakšsignālu dekompozīcija

Variāciju apakšsignālu dekompozīcijas (*VMD*) metode ir adaptīvs signāla

dekompozīcijas modelis, kurā tiek iegūts diskrēts skaits šī signāla kvazi-ortogonālu apakšsignālu, kur šo apakšsignālu summa veido oriģinālo signālu [118]. Šis modelis ir klasiskā Vīnera filtra [43], ko izmanto signālu trokšņu noņemšanai, vispārinājums. Variāciju apakšsignālu dekompozīcijas modelis tiek efektīvi optimizēts, izmantojot reizinātāju mainīgo virzienu metodi (*Alternate Direction Method of Multipliers, ADMM*). Variāciju apakšsignālu dekompozīcijas metode balstās uz trim metodēm – Furjē transformāciju (skat. 2.1.1. punktu), Hilberta transformāciju (skat. 2.1.2. punktu) un Vīnera filtrāciju (skat. 2.1.3. punktu).

2.1.1. Furjē transformācija

Furjē transformācija ir spektrālās analīzes metode, kas transformē signālu no laika apgabala uz frekvenču apgabalu, veicot oriģinālā signāla sadalīšanu uz sinusoidālām funkcijām, ko sauc par harmonikām [48]. Furjē transformācijas rezultātā iegūtais frekvenču spektrs reprezentē šī signāla amplitūdu un fāžu sadalījumu frekvenču apgabalā. Diskrētā signāla $f = \{f(t), t = 0,...,N-1\}$ gadījumā izmanto diskrēto Furjē transformāciju [48], kur tās vērtību $F(n)$ frekvences solī n aprēķina atbilstoši izteiksmei (2.1):

$$F(n) = \sum_{t=0}^{N-1} f(t) e^{-j\frac{2\pi}{N}nt}, \tag{2.1.}$$

kur N – diskrētā signāla elementu skaits;

t – laika soļu indekss, $t = 0,\ldots,N-1$;

$j^2 = -1$ – imaginārais lielums;

n – frekvences soļu indekss, $n = 0,\ldots,N-1$;

π – matemātiskā konstante, $\pi = 3.1415$.

Attiecīgi inversās diskrētās Furjē transformācijas vērtību $f(t)$ laika solī t var iegūt kā izteiksmi (2.2):

$$f(t) = \frac{1}{N} \sum_{n=0}^{N-1} F(n) e^{j\frac{2\pi}{N}nt}. \tag{2.2.}$$

Diskrētās Furjē transformācijas vērtības ir sinusoidālu signālu kompleksas amplitūdas.

2.1.2. *Hilberta transformācija*

Hilberta transformācija [118] ir lineārs operators, kas katru funkciju $f(t)$ kartē uz atbilstošu funkciju $Hf(t)$. Hilberta transformācija ir fāžu filtrs, kas izlaiž cauri visas signāla frekvences ω, bet izmaina signāla frekvenču fāzes, nobīdot negatīvo frekvenču fāzes par $\pi/2$ radiāniem, bet pozitīvo frekvenču fāzes par $-\pi/2$ radiāniem. Hilberta transformāciju apraksta pārraides funkcija $\hat{h}(\omega)$ [118], ko var izteikt ar izteiksmi (2.3.):

$$\hat{h}(\omega) = -j * sgn(\omega) = -\frac{j\omega}{|\omega|} = \begin{cases} e^{\frac{j\pi}{2}}, & \text{ja} \quad \omega < 0; \\ 0, & \text{ja} \quad \omega = 0; \\ e^{\frac{-j\pi}{2}}, & \text{ja} \quad \omega > 0, \end{cases} \tag{2.3.}$$

kur $*$ – konvolūcijas zīme.

Izteiksmē (2.3.) $sgn(\omega)$ noteikts pēc formulas (2.4.):

$$sgn(\omega) = \begin{cases} -1, & \text{ja} \quad \omega < 0; \\ 0, & \text{ja} \quad \omega = 0; \\ 1, & \text{ja} \quad \omega > 0. \end{cases} \tag{2.4.}$$

Pārraides funkcija ir impulsa reakcijas funkcijas Furjē transformācija un pārraides funkcijai (2.3) atbilstošo impulsa reakcijas funkciju laika apgabalā – Dīraka delta funkciju [118] apraksta izteiksme (2.5):

$$h(t) = \frac{1}{\pi t}. \tag{2.5.}$$

Inversā Hilberta transformācija [118] tiek iegūta kā $-Hf(t)$.

2.1.3. *Gausa regularizācija un Vīnera filtrs*

Bieži oriģinālo, tīro signālu ir degradējis aditīvais, baltais Gausa troksnis, kas ir gadījuma signāls ar vidējo vērtību nulle, standartnovirzi viens un konstantu amplitūdu frekvenču apgabalā [118]. Tad novēroto signālu $f_o(t)$ var izteikt kā oriģinālā signāla f un balta Gausa trokšņa η summu (2.6 izteiksme):

$$f_o = f + \eta. \tag{2.6.}$$

Uzdevums ir restaurēt nezināmo, oriģinālo signālu f. Ja oriģinālais signāls mainās vienmērīgi, tad, izmantojot minimizācijas problēmu ar Tihonova regularizāciju [118], var iegūt no trokšņiem attīrītu signālu (2.7. izteiksme):

$$\min_f \left\{ f - f_{o2}^2 + \alpha \partial_t f_2^2 \right\}, \tag{2.7.}$$

kur α ir Gausa (baltā) trokšņa dispersija un $d_t f$ ir nezināmās funkcijas f atvasinājums pēc laika t. Vienādojums (2.7.) ir Gausa regularizētā vidējo kvadrātu minimuma problēma [118], no kuras var iegūt Eilera-Lagranža vienādojumus (2.8):

$$f - f_o = \alpha \partial_t^2 f. \tag{2.8.}$$

Šos (2.8) Eilera-Lagranža vienādojumus [118] risina spektrālajā (frekvenču) apgabalā, atbilstoši izteiksmei (2.9.):

$$\hat{f}(\omega) = \frac{\hat{f}_o(\omega)}{1 + \alpha \omega^2}, \tag{2.9.}$$

kur $\hat{f}(\omega)$ – no trokšņiem attīrīta signāla $f(t)$ Furjē transformācija;

$\hat{f}_o(\omega)$ – novērota, trokšņaina signāla $f_o(t)$ Furjē transformācija.

Restaurētais signāls $f(t)$ ir zemo frekvenču (apmēram pie $\omega = 0$) izlase [118] no novērotā signāla $f_o(t)$.

2.1.4. Variāciju apakšsignālu dekompozīcijas algoritms

Variāciju apakšsignālu dekompozīcijas modelī tiek pieņemts, ka reālo vērtību signāls f sastāv no vairākiem apakšsignāliem u_k, kur $k = 1, \ldots, K$. Apakšsignāls jeb iekšējo svārstību funkcija ir amplitūdas un frekvences modulētais (AM-FM) signāls [118], un to var pierakstīt atbilstoši (2.10.) izteiksmei:

$$u_k(t) = A_k(t) \cos\left(\phi_k(t)\right), \tag{2.10.}$$

kur k – apakšsignāla numurs, $k = 1, \ldots, K$, kur K ir apakšsignālu skaits;

$A_k(t)$ – k-tā apakšsignāla amplitūda;

$\phi_k(t)$ – k-tā apakšsignāla fāze;

t – laiks.

Katram apakšsignālam frekvences ω mainās nelielā diapazonā. Katram no K apakšsignāliem ir jābūt centrētam ap centra frekvenci ω_k, kas tiek aprēķināta dekompozīcijas laikā. Uzdevums ir minimizēt K apakšsignālu frekvenču joslu platumu summu ar nosacījumu, ka visu K apakšsignālu summa ir vienāda ar oriģinālo signālu, un, lai aprēķinātu katra apakšsignāla frekvenču joslas platumu, tiek izmantots algoritms, kas sastāv no trīs soļiem [118].

1. Katram apakšsignālam $u_k(t)$ tiek aprēķināta Hilberta transformācija $u_k^H(t)$, kas tiek izmantota, lai iegūtu signāla $u_k(t)$ analītisko funkciju $f_A(t)$ (2.11.):

$$f_A(t) = u_k(t) + ju_k^H(t) = A(t)e^{j\phi(t)} = \left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t), \qquad (2.11.)$$

kur j – imaginārā vienība, $j^2 = -1$;
$\delta(t)$ – Dīraka delta funkcija;
$*$ – konvolūcijas simbols.

Tas nozīmē, ka reālais signāls $u_k(t)$ tiek paplašināts kompleksajā plaknē, iegūstot šī signāla harmonisko konjugātu [76]. Tālāk tiek iegūts šīs analītiskās funkcijas vienpusējs frekvenču spektrs, kas satur tikai pozitīvas frekvences, un šis spektrs ir centrēts pie frekvences ω_k.

2. Katram apakšsignālam atbilstošā analītiskā funkcija tiek reizināta ar eksponenciālo funkciju $e^{-j\omega_k t}$, atbilstoši formulai (2.12.):

$$\left[u_k(t) + ju_k^H(t)\right]e^{-j\omega_k t}. \qquad (2.12.)$$

Tādā veidā tiek veikta signāla demodulācija – apakšsignāla frekvenču spektrs tiek nobīdīts uz pamatjoslu jeb signāla oriģinālo frekvenču diapazonu.

3. Iekšējo svārstību funkcijas $u_k(t)$ frekvenču joslas platums tagad tiek aprēķināts, izmantojot šī demodulētā signāla gradienta L^2-normas (Eiklīda normas) kvadrātu, un, izmantojot variāciju aprēķinu formulējumu, tiek iegūta šāda minimizācijas problēma (2.13.):

$$\min_{u_k,\omega_k}\left\{\sum_{k=1}^{K}\left\|\partial t\left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\} \qquad (2.13.)$$

ar nosacījumu $\sum_{k=1}^{K} u_k(t) = f(t)$,

kur f – oriģinālais signāls;

∂t – atvasinājums pēc laika;

$\| \|_2$ – L^2-norma;

K – apakšsignālu skaits.

Šī (2.13) problēma ir šāda: ir nepieciešams minimizēt K apakšsignālu frekvenču joslu platumu summu ar nosacījumu, ka šo K apakšsignālu summa ir vienāda ar oriģinālo signālu. Tātad nezināmie ir K apakšsignālu centra frekvences un K iekšējo svārstību funkcijas, kas ir centrētas pie šīm frekvencēm. Daži no nezināmajiem ir funkcijas, tāpēc tiek izmantoti variāciju aprēķini.

Minimizācijas problēmas (2.13.) mērķfunkcija var tikt izteikta kā beznosacījumu optimizācijas problēma, izmantojot paplašināto Lagranža funkciju [118]. Atrisinājums minimizācijas problēmai (2.13.) tiek atrasts, secīgi izmantojot iteratīvu optimizācijas metodi, ko sauc par reizinātāju mainīgo virzienu metodi (*ADMM*). Paplašināto Lagranža funkcija $\mathcal{L}$ ar soda parametru $\alpha > 0$ var parādīt izteiksmes (2.14.) veidā:

$$\mathcal{L}\left(u_k, \omega_k, \lambda\right) = \alpha \sum_{k=1}^{K} \left\| \partial t \left[\left(\delta\left(t\right) + \frac{j}{\pi t} \right) * u_k\left(t\right) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle, \tag{2.14.}$$

kur $\langle .,. \rangle$ – skalārs reizinājums;

λ – Lagranža reizinātājs.

Izmantojot reizinātāju mainīgo virzienu metodi, atrisinājums paplašinātajai Lagranža funkcijai (2.14.) tiek atrasts vienam mainīgajam, pieņemot, ka visi pārējie mainīgie ir zināmi [118]. Ja α ir liels, tad centra frekvences ω_k var tikt precīzi noteiktas, bet tiek stingri pārkāpti optimizācijas nosacījumi, savukārt priekš maziem α optimizācijas nosacījumi tiek izpildīti vai minimāli pārkāpti, bet optimizācija paliek skaitliski slikti nosacīta [12]. Lai mazinātu šo problēmu starp precīzu centra frekvenču noteikšanu un precīzu oriģinālā signāla rekonstrukciju, tiek ieviests Lagranža reizinātājs λ, un izmantojot paplašināto Lagranža funkciju tiek mēģināts atrast tādas nezināmo vērtības, kas ļauj nepārkāpt vai minimāli pārkāpt optimizācijas nosacījumus, bet soda parametram α nav jābūt ļoti mazam [12].

Lai iegūtu korektus rezultātus arī abos signāla galapunktos, izmantojot variāciju

apakšsignālu dekompozīcijas algoritmu, oriģinālais signāls tiek paplašināts, izmantojot "spoguļošanu" [118]. Paplašinātais signāls ir divas reizes garāks par oriģinālo signālu, un tā garums ir T, tāpat tiek pieņemts, ka paplašinātā signāla frekvences ω mainās robežās [–0,5; 0,5], tādējādi iegūstot normalizēto frekvenču sadalījumu [119].

Katrā iterācijas solī tiek aprēķināts atlikums, kuru iegūst, izmantojot iepriekšējā un pašreizējā iterācijā iegūtās apakšsignālu vērtības [119]. Pirmajā iterācijas solī atlikums $S_k^{n-1} = 0$. Iterācijas solī n pirmajam apakšsignālam ($k = 1$) atbilstošais atlikums S_k^n tiek aprēķināts pēc izteiksmes (2.15.):

$$S_k^n = \hat{u}_K^n + S_K^{n-1} - \hat{u}_1^n, \tag{2.15.}$$

kur $\hat{u}_K^n$ – *K-tā* apakšsignāla u_K Furjē transformācija iterācijas solī n;

S_K^{n-1} – iepriekšējā iterācijas solī $n-1$ pēdējais iegūtais atlikums, kas atbilst *K-tā* apakšsignāla atlikumam;

$\hat{u}_1^n$ – pirmā apakšsignāla u_1 Furjē transformācija iterācijas solī n;

K – apakšsignālu skaits.

Pārējiem apakšsignāliem (ar indeksu $k = 2, \ldots, K$) atbilstošais pašreizējais atlikums n iterācijas solī S_k^n tiek aprēķināts pēc izteiksmes (2.16.):

$$S_k^n = \hat{u}_{k-1}^{n+1} + S_{k-1}^n - \hat{u}_k^n, \tag{2.16.}$$

kur $\hat{u}_{k-1}^{n+1}$ – $k-1$ apakšsignāla u_{k-1} Furjē transformācija iterācijas solī $n+1$;

S_{k-1}^n – iepriekšējam apakšsignālam (ar indeksu $k-1$) atbilstošais atlikums iterācijas solī n;

$\hat{u}_k^n$ – *k-tā* apakšsignāla u_k Furjē transformācija iterācijas solī n.

Katrā iterācijas solī $n+1$, *k-tā* apakšsignāla u_k Furjē transformācija $\hat{u}_k^{n+1}$ tiek aprēķināta, izmantojot izteiksmi (2.17.), kas ir pašreizējā atlikuma S_k^n filtrācija, izmantojot Vīnera filtru [119]:

$$\hat{u}_k^{n+1} = \frac{\hat{f} - S_k^n - \dfrac{\lambda^n}{2}}{1 + \alpha(\omega - \omega_k^n)^2}, \tag{2.17.}$$

kur $\hat{f}$ – paplašinātā signāla Furjē transformācija;

S_k^n – *k-tām* apakšsignālam atbilstošais pašreizējais atlikums iterācijas solī n;

$\hat{\lambda}^n$ – Lagranža reizinātāja Furjē transformācija iterācijas solī n;

α – Gausa (baltā) trokšņa dispersija;

$\frac{1}{(\omega-\omega_k^n)^2}$ – frekvenču spektrs, kur ω ir paplašināta signāla f frekvences un ω_k^n ir *k-tā* apakšsignāla centra frekvence iterācijas solī n.

Centra frekvenci ω_k^{n+1} [119] $n+1$ iterācijas solī *k-tām* apakšsignālam aprēķina pēc izteiksmes (2.18.):

$$\omega_k^{n+1} = \frac{\sum_{i=\frac{T}{2}+1}^{T} \omega(i) \left|\hat{u}(i)_k^{n+1}\right|^2}{\sum_{i=\frac{T}{2}+1}^{T} \left|\hat{u}(i)_k^{n+1}\right|^2}, \tag{2.18.}$$

kur T – frekvenču elementu indekss;

ω – paplašināta signāla f frekvences.

Lagranža reizinātājs $\hat{\lambda}^{n+1}$ [119] iterācijas solī $n+1$ tiek atjaunots pēc izteiksmes (2.19.):

$$\hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau\left(\sum_{k=1}^{K} \hat{u}_k^{n+1} - \hat{f}\right), \tag{2.19.}$$

kur τ – laika solis divu mainīgo gradienta pacelšanas metodei.

Iterācijas turpinās, kamēr netiek sasniegts maksimālais iteraciju skaits, kur *VMD* metodes autori piedāvā 500 iterācijas, vai kamēr neizpildās konverģences nosacījums [119], ko aprēķina atbilstoši izteiksmei (2.20.):

$$\left|\sum_{k=1}^{K} 1/T\left[\hat{u}_k^n - \hat{u}_k^{n-1}\right]\left[\overline{\hat{u}_k^n - \hat{u}_k^{n-1}}\right]^T\right| < \varepsilon, \tag{2.20.}$$

kur T – paplašinātā signāla garums;

$\left[\overline{\hat{u}_k^n - \hat{u}_k^{n-1}}\right]^T$ – *k-tā* apakšsignāla u_k *n*-tajā un *n-1* iterācijā vērtību starpības kompleksā konjugāta transponētais vektors;

ε – konverģences kritērijs;

$|\ \ |$ – absolūtā vērtība.

Kad visi apakšsignāli frekvenču apgabalā ir iegūti, tad, izmantojot inverso

diskrēto Furjē transformāciju (2.2.), tiek iegūti apakšsignāli laika apgabalā. Šie apakšsignāli atbilst paplašinātajam signālam. Izņemot ārā no apakšsignāliem tās signāla daļas, kas tika pievienotas oriģinālajam signālam "spoguļošanas" rezultātā, tiek iegūtas oriģinālā signāla iekšējo svārstību funkcijas.

2.2. Fāžu telpas rekonstrukcija ar laika aiztures metodi

Haotiskās, dinamiskās sistēmas ir bieži sastopamas, un lielai daļai nav pieejami vienādojumi, kas aprakstītu šo sistēmu. Līdz ar to vienīgais veids, kā saprast šādu sistēmu, ir izmantot pieejamo laika rindu. Sistēmu, kuru stāvoklis mainās laikā, sauc par dinamisku sistēmu [53]. Novērotā laika rinda, principā, tiek iegūta no kādas dinamiskas sistēmas ar vai bez trokšņa ietekmes [58].

Fāžu telpa matemātikā un fizikā ir abstrakta telpa, kura attēlo visu sistēmas iespējamo stāvokļu kopu, kur katram iespējamajam sistēmas stāvoklim atbilst punkts fāžu telpā [11]. Tādējādi sistēma tiek aizvietota ar šīs sistēmas iespējamo stāvokļu kopumu. Fāžu telpa tiek izmantota, lai attēlotu sarežģītas sistēmas stāvokli ar vienu vienīgu punktu, bet šīs sistēmas evolūciju – ar šī punkta pārvietošanos pa fāžu trajektoriju, ko sauc par orbītu [49].

Deterministiskā sistēma ir sistēma, kuru reglamentē precīzi likumi un kura neizmanto gadījuma elementus savu nākotnes stāvokļu attīstībā [39]. Deterministiskas, dinamiskas sistēmas stāvoklis laika solī t satur visu nepieciešamo informāciju, lai noteiktu sistēmas nākotnes stāvokļus. Haotiska sistēma matemātiski tiek definēta kā dinamiska sistēma vismaz ar vienu pozitīvu Ļapunova eksponentes pakāpes rādītāju jeb, citiem vārdiem sakot, haotiska sistēma ir dinamiska sistēma, kas ir jutīga pret sākuma nosacījumiem – sistēmas mainīgo vērtībām sākuma laikā [101]. Haotiskas sistēmas ir deterministiskas un nelineāras un ir prognozējamas īstermiņā, bet ilgtermiņa prognozēšana var būt sarežģīta, vai pat neiespējama [31].

Haotiskām, dinamiskām sistēmām fāžu telpa un sistēmas matemātiskais apraksts bieži vien nav zināms. Tomēr ir iespējams iegūt rekonstruēto fāžu telpu, kura topoloģiski ir ekvivalenta oriģinālajai sistēmai, un ir spējīga atjaunot sistēmas dinamiku [57]. Fāžu telpas rekonstrukcija ir pirmais solis nelineāru laika rindu analīzē un to prognozēšanā [115].

Haotiskām laika rindām izmanto fāžu telpas rekonstrukciju ar laika aiztures metodi pēc Takensa teorēmas [49]. Izmantojot Takensa teorēmu, haotisku un

dinamisku sistēmu var attēlot ar aiztures vektoru palīdzību. Ja dota laika rinda y ar garumu N (skat. izteiksmi (2.21.)):

$$y = \{y(1), y(2), \ldots, y(N)\}, \tag{2.21.}$$

tad tās i-*to* aiztures vektoru var uzrakstīt kā izteiksmi (2.22.):

$$y_i = \left[y(t_i), y(t_i + \tau), y(t_i + 2\tau), \ldots, y(t_i + (m-1)\tau) \right], \tag{2.22.}$$

kur i – aiztures vektoru skaits, $i = 1, \ldots, N - (m-1)\tau$;

m – rekonstruētās fāžu telpas dimensija;

τ – laika aizture.

Aiztures vektoru kopa veido rekonstruēto fāžu telpu [49], ko apraksta matrica (2.23.):

$$PhS = \begin{bmatrix} y(1) & y(1+\tau) & y(1+2\tau) & \ldots & y(1+(m\text{-}1)\tau) \\ y(2) & y(2+\tau) & y(2+2\tau) & \ldots & y(2+(m\text{-}1)\tau) \\ \ldots & \ldots & \ldots & & \ldots \\ y(M) & y(M+\tau) & y(M+2\tau) & \ldots & y(M+(m\text{-}1)\tau) \end{bmatrix}, \tag{2.23.}$$

kur M – punktu skaits fāžu telpā, $M = N - (m+1)\tau$.

Izmantojot fāžu telpas rekonstrukcijas metodi ar laika aizturi, ir nepieciešams atrast divu parametru – iegultās dimensijas m un laika aiztures τ piemērotās vērtības. No šiem parametriem ir atkarīga laika rindu prognozēšanas precizitāte. Prognozēšanas uzdevuma formālajā nostādnē, meklējot funkcionālo sakarību (1.7.), kā loga garums m sākotnēji, pirms tiek pabeigta datu pirmapstrāde, tiek izmantota rekonstruētās fāžu telpas dimensija.

2.3. Pakāpeniskā regresijas analīze

Pakāpeniskā regresijas analīze ir sistemātiska metode secīgai pazīmju atlasīšanai, kur pazīmes jeb atribūti tiek pievienoti multilineāram modelim un izņemti no tā automātisko procedūru rezultātā, pamatojoties uz atribūtu statistisko nozīmi regresijas analīzē [98]. Pakāpeniskās regresijas analīzes algoritms sastāv no trim soļiem [16].

1. Sākotnējais pakāpeniskās regresijas modelis neietver nevienu atribūtu.

2. Katrā solī tiek aprēķināta F statistikas p vērtība, lai testētu modeļus ar noteiktiem atribūtiem [111]. F statistika tiek aprēķināta pēc izteiksmes (2.24.):

$$F = \frac{\left(\sum_{i=1}^{N}\left(y_i - \hat{y}_{1i}\right)^2 - \sum_{i=1}^{N}\left(y_i - \hat{y}_{2i}\right)^2\right)\left(N - p_2\right)}{\left(p_2 - p_1\right)\left(\sum_{i=1}^{N}\left(y_i - \hat{y}_{2i}\right)^2\right)}, \qquad (2.24.)$$

kur y – novērotais, atkarīgais parametrs;

$\hat{y}_1$ – pirmā modeļa prognozētais parametrs;

$\hat{y}_2$ – otrā modeļa prognozētais parametrs;

p_1 – pirmā modeļa atribūtu skaits;

p_2 – otrā modeļa atribūtu skaits;

N – novērojumu skaits.

Prognozējamais parametrs $\hat{y}_m$ m-tajā multilineārajā modelī tiek iegūts pēc vienādojuma (2.25.):

$$\hat{y}_m = b_{m0} + b_{m1}x_{m1} + \ldots + b_{mp_m}x_{mp_m}, \qquad (2.25.)$$

kur b_{m0} – *m-tā* modeļa regresijas konstante;

b_{m1} – *m-tā* modeļa pirmā atribūta koeficients;

b_{mp_m} – *m-tā* modeļa p_m atribūta koeficients;

x_{m1} – *m-tā* modeļa pirmais atribūts;

x_{mp_m} – *m-tā* modeļa p_m-*tais* atribūts.

Jebkuram atribūtam, kurš pašlaik nav pievienots modelim un kuram p vērtība ir mazāka par noteikto pievienošanas slieksni, tiek izvirzīta nulles hipotēze, ka, pievienojot šo atribūtu modelim, tam varētu būt nulles koeficients. Ja ir pietiekami pierādījumi, lai noraidītu nulles hipotēzi, tad atribūts ar mazāko p vērtību tiek pievienots modelim. Šis solis tiek atkārtots, kamēr pastāv atribūti, kuri nav pievienoti modelim, un kuriem p vērtība ir mazāka par noteikto pievienošanas slieksni. Ja šādu atribūtu nav, tad notiek pāreja uz trešo soli.

3. Jebkuram atribūtam, kas jau ir pievienots modelim un kuram p vērtība pārsniedz izņemšanas slieksni, tiek izvirzīta nulles hipotēze, ka šim atribūtam ir nulles koeficients. Ja nav pietiekamu pierādījumu, lai noraidītu hipotēzi, atribūts ar

lielāko p vērtību tiek izņemts no modeļa, un notiek atgriešanās otrajā solī.

Pretējā gadījumā, ja vairs netiek iegūts labāks modelis, t. i., kad visi nozīmīgie atribūti jau ir iekļauti modelī, process apstājas. Informatīvāko atribūtu indeksi (atribūtu kolonnu numuri atribūtu matricā) tiek saglabāti. Tomēr pakāpeniskās regresijas analīze nav globāli optimāla, un iegūtais regresijas modelis var būt tikai lokāli optimāls [16].

2.4. Principiālo komponentu analīze

Principiālo komponentu analīze (*PCA*) ir statistiska pazīmju izvilkšanas metode, kas izmanto ortogonālu transformāciju, lai pārveidotu iespējami korelējošo datu kopu uz lineāri nekorelējošu datu kopu, kur šīs datu kopas pazīmes sauc par principiālām komponentēm [79]. *PCA* algoritms sastāv no lineārās transformācijas $\hat{x}_1, \hat{x}_2, \hat{x}_3, \ldots, \hat{x}_p$ atrašanas oriģinālo pazīmju (atribūtu) kopai $x_1, x_2, x_3, \ldots, x_m$ [79]. Šeit m ir sākotnējās atribūtu kopas dimensiju skaits, bet p ir principiālo komponentu skaits. Tādējādi *PCA* metode arī var tikt izmantota atribūtu kopas dimensiju skaita reducēšanai. Principiālās komponentes $\hat{x}$ tiek izvēlētas tādā veidā, ka pirmā principiālā komponente $\hat{x}_1$ satur visvairāk dispersijas no oriģinālajiem datiem, otrā principiālā komponente $\hat{x}_2$ satur otro lielāko dispersijas lielumu no oriģinālajiem datiem, turklāt tā lineāri nekorelē ar $\hat{x}_1$, utt. [79]. Tātad *PCA* metodes mērķis ir atrast ortogonālu komponentu kopu.

Pirmais solis *PCA* algoritmā ir normalizēt datus tā, lai katram atribūtam vidējā vērtība būtu nulle. Tad tiek izmantota ortogonalizācijas metode, lai aprēķinātu principiālās komponentes no normalizētiem datiem. Paraugkopas kovariācijas matricas C elementu C_{jk} [96] starp atribūtu kopas x *j-to* un *k-to* atribūtu iegūst pēc izteiksmes (2.26.):

$$C_{jk} = \frac{1}{N-1} \sum_{i=1}^{N} \left(x_{ij} - \bar{x}_j \right) \left(x_{ik} - \bar{x}_k \right), \tag{2.26.}$$

kur x_{ij} – atribūtu kopas x *j-tā* atribūta *i-tais* elements;

x_{ik} – atribūtu kopas x k-*tā* atribūta *i-tais* elements;

$\bar{x}_j$ – atribūtu kopas x *j-tā* atribūta vidējā vērtība;

$\bar{x}_k$ – atribūtu kopas x *k-tā* atribūta vidējā vērtība;

N – atribūtu novērojumu skaits.

Veicot paraugkopas kovariācijas matricas C īpašvērtību dekompozīciju, tiek iegūta īpašvektoru kopa M un īpašvērtības l. Principiālo komponentu kopas i*-tā* komponente $\hat{x}_i$ [96] tiek iegūta pēc izteiksmes (2.27.):

$$\hat{x}_i = Mx_i, \ i \in 1,...,m. \tag{2.27.}$$

Paraugkopas kovariācijas matricas reizinājums ar i*-to* principiālo komponenti (transformēto atribūtu) ir vienāds ar i*-tās* īpašvērtības un i*-tā* transformētā atribūta reizinājumu, tādēļ ir spēkā izteiksme (2.28.):

$$C\hat{x}_i = l_i \hat{x}_i, \ i \in 1,\ldots,m, \tag{2.28.}$$

kur l_i ir i*-tā* pēc kārtas lielākā paraugkopas kovariācijas matricas C īpašvērtība [96].

2.5. Mākslīgie neironu tīkli

Mākslīgie neironu tīkli (*Artificial Neural Network, ANN*), ir mākslīgā intelekta paveids, kas mēģina imitēt bioloģisko neironu funkciju, pēc kuras strādā cilvēka smadzenes [85]. *ANN* ir viens no precīzākajiem prognozēšanas modeļiem, kuru struktūra padara tos piemērotus prognozēšanas uzdevumiem ar labu precizitāti.

Mākslīgie neironu tīkli mācās no piemēriem un var atrast funkcionālas attiecības starp datiem, pat, ja šīs attiecības nav zināmas, vai to fizikālā nozīme ir neskaidra [70]. Tādējādi *ANN* ir labi piemēroti problēmām, kuru risinājumam ir nepieciešamas padziļinātas zināšanas, bet kurām ir pieejami pietiekami daudz datu.

Mākslīgie neironu tīkli var vispārināt likumsakarības starp datos esošajiem tēliem (*patterns*). Pēc apmācības no ieejas datiem, *ANN* bieži vien var korekti apstrādāt agrāk neredzētos datu ierakstus, pat, ja tie satur troksni [113]. Tie var daudz labāk par lielāko daļu citu metožu prognozēt trokšņainas, nelineāras, nestacionāras un haotiskas laika rindas [85]. Mākslīgie neironu tīkli arī ir universālais funkciju aproksimators. Ir pierādīts, ka mākslīgie neironu tīkli var aproksimēt jebkuru nepārtrauktu funkciju ar jebkuru precizitāti [112].

2.5.1. *Daudzslāņu perceptrons*

Mākslīgo neironu tīklu struktūra iekļauj ieejas datus un mākslīgos neironus. Lai prognozētu laika rindas vērtību $y(t+1)$, *ANN* veic nezināmās funkcijas (2.29.) modelēšanu, kur šīs funkcijas parametri ir laika rindas vēsturiskās vērtības [112]:

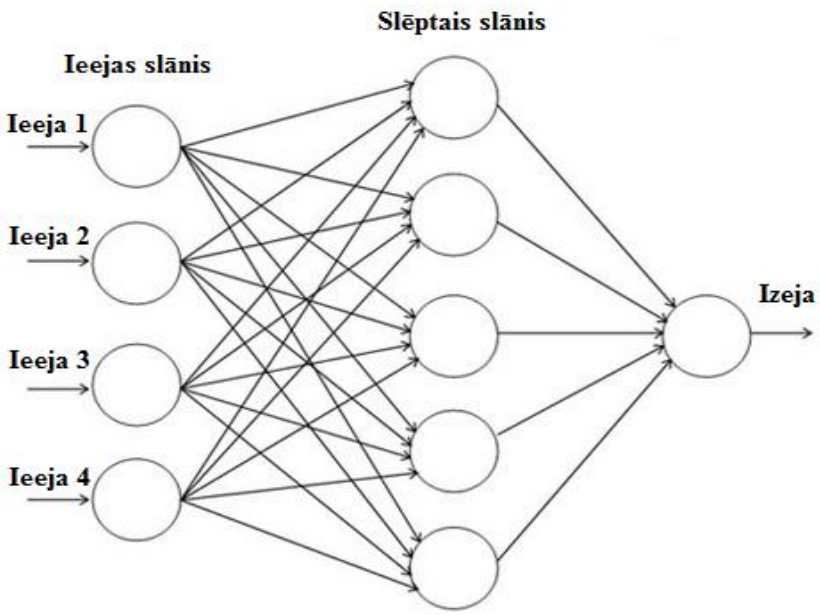
$$y(t+1) = f\left(y(t), y(t-1), \ldots, y(t-m+1)\right) + \varepsilon(t+1), \tag{2.29.}$$

kur $y(t)$ – novērojums laika solī t;

m – loga garums;

$\varepsilon(t+1)$ – baltais troksnis laika solī $t+1$.

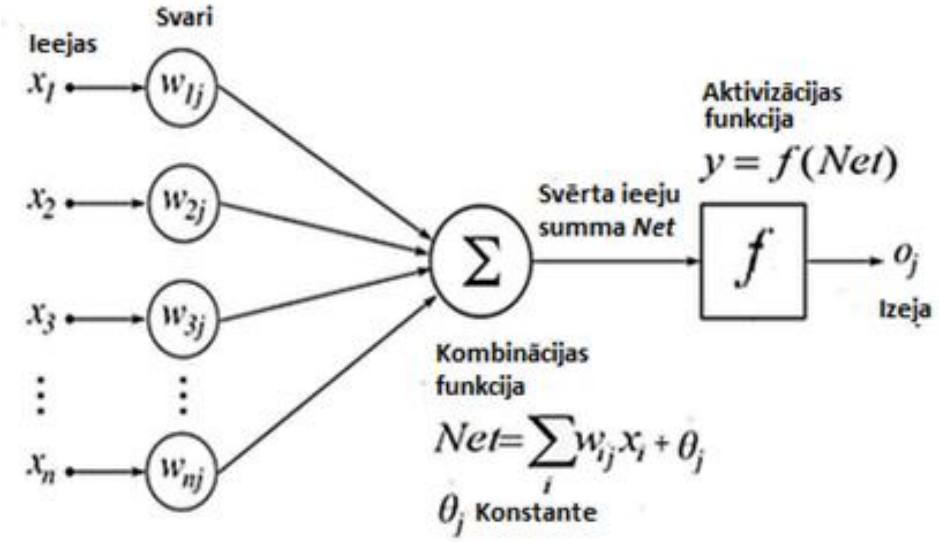
Viens no izplatītākajiem mākslīgo neironu tīklu modeļiem ir daudzslāņu perceptrons (*Multilayer Perceptron*). Daudzslāņu perceptrons (skat. 2.1. att.) iekļauj ieejas slāni, izejas slāni un vienu vai vairākus starpslāņus, ko sauc par slēptajiem slāņiem [85]. Promocijas darbā, risinot laika rindu prognozēšanas uzdevumu, tiek izmantota mākslīgo neironu tīklu topoloģija, kur izejas slānis satur vienu neironu, kas izejā dod neironu tīkla prognozēto vērtību, bet tīkla ieeju skaits var būt dažāds. Topoloģijas attēlā (skat. 2.1. att.) ir parādīts piemērs ar četrām ieejām.



2.1. att. Daudzslāņu perceptrona topoloģija.

Skalārie svari kopā ar tīkla arhitektūru glabā uztrenēta tīkla zināšanas un nosaka savienojumu stiprumu starp savā starpā savienotiem neironiem [90]. Katrs neirons

saņem svērtās ieejas no iepriekšējā slāņa, tās tiek summētas, izmantojot kombināciju funkciju, un šīs summēšanas rezultāts tiek padots kā arguments aktivizācijas funkcijai. Aktivizācijas funkcijas vērtība ir neirona izeja (skat. 2.2. att.), un tā tiek padota uz nākamo slāni [85].



2.2. att. Informācijas izplatīšanas neironu tīklā [50].

Dažas aktivizācijas funkcijas, ko izmanto mākslīgie neironu tīkli [46], un, kas arī tiek izmantotas šajā darbā, ir dotas 2.1. tabulā.

2.1. tabula

Mākslīgo neironu tīklu aktivizācijas funkcijas

Pārraides funkcijas	
Nosaukums	**Formula**
Lineāra	$y(x) = x$
Sigmoidāla (loģistiskā)	$y(x) = \frac{1}{1+e^{-x}}$
Hiperboliskais tangenss	$y(x) = \frac{e^{2x}-1}{e^{2x}+1}$

Aktivizācijas funkcija normalizē tās argumentu jeb ieejas datus uz kādu noteiktu intervālu. Piemēram, sigmoidāla aktivizācijas funkcija normalizē tās ieejas datus uz intervālu [0; 1], kamēr hiperboliskā tangensa aktivizācijas funkcija uz intervālu [–1; 1]. Ja neironu tīkla ieejas dati satur tikai pozitīvas vērtības, tad slēptā slāņa neironiem kā aktivizācijas funkcija var tikt izmantota sigmoidāla funkcija, bet, ja ieejas

dati var saturēt arī negatīvās vērtības, tad savukārt slēptā slāņa neironiem var tikt izmantota hiperboliskā tangensa funkcija. Lineārā aktivizācijas funkcija izejā dod to pašu vērtību, ko saņem ieejā, tāpēc to bieži izmanto kā izejas slāņa neirona aktivizācijas funkciju, jo neironu tīklā aprēķinātās laika rindas prognozes var tikt kropļotas, ja izejas slānī tiek izmantota kāda cita aktivizācijas funkcija [46].

Tādējādi, izmantojot mākslīgo neironu tīklu ar vienu slēpto slāni, vienu neironu izejas slānī ar lineāro aktivizācijas funkciju, sakarība starp prognozējamo laika rindas vērtību $\hat{y}(t+1)$ un neironu tīkla ieejas datiem: laika rindas novērojumiem $(y(t),...,y(t-m+1))$ ir izsakāma ar izteiksmi (2.30.):

$$\hat{y}(t+1) = w_0 + \sum_{j=1}^{q} w_j f\left(w_{0,j} + \sum_{i=1}^{p} w_{i,j} y(t-i+1) \right) + \varepsilon(t+1), \qquad (2.30.)$$

kur w_o – izejas slāņa konstante (*bias*);

q – slēptā slāņa neironu skaits;

w_j – svars starp slēptā slāņa *j*-to neironu un izejas slāņa neironu;

f – slēptā slāņa neironu aktivizācijas funkcija;

$w_{o,j}$ – slēptā slāņa *j*-tā konstante (*bias*);

$w_{i,j}$ – svars starp ieejas slāņa i-to neironu un slēptā slāņa *j*-to neironu;

p – ieejas slāņa neironu skaits.

Neironu tīkla apmācība ir beznosacījumu nelineāras optimizācijas problēma. Pastāv vairāki neironu tīkla apmācības algoritmi. Visizplatītākais ir gradienta nolaišanās kļūdas atgriezeniskās izplatīšanās (*Gradient Descent Backpropagation*) algoritms, kas koriģē svarus un konstantes gradienta nolaišanās negatīvajā virzienā (gradienta minimuma virzienā), tomēr šī algoritma galvenā problēma ir lēnā konverģence, kā arī optimizācija var iestrēgt lokālajā minimumā [85]. Lai risinātu šīs problēmas, tiek izmantoti citi neironu tīkla apmācības algoritmi.

Viens no tiem ir Levenberga-Markvardta kļūdas atgriezeniskās izplatīšanās algoritms ar Beijesa regularizāciju (*Levenberg-Marquardt Backpropagation Algorithm with Bayesian Regularization*). Tas atjauno neironu tīkla modeļa svaru un konstantu vērtības atbilstoši vienai no populārākajām un efektīvākajām nelineārās optimizācijas metodēm – Levenberga-Markvardta optimizācijas algoritmam.

Levenberga-Markvardta optimizācijas algoritms ir Ņūtona metodes alternatīva,

kas meklē risinājumu, izmantojot nelineāru mazāko kvadrātu optimizāciju, un šai optimizācijas metodei piemīt ātrāka konverģence, un tā var atrast labāku lokālo minimumu, salīdzinot ar gradienta nolaišanās metodi [85].

Neironu tīkla apmācības mērķis ir samazināt globālo kļūdu, kura tiek aprēķināta, izmantojot kādu noteiktu mērķfunkciju. Levenberga-Markvardta kļūdas atgriezeniskās izplatīšanās algoritmam ar Beijesa regularizāciju kā mērķfunkcija tiek izmantota regularizēta vidējās kvadrātiskās kļūdas zaudējumu funkcija MSE_{reg} [23], ko apraksta ar izteiksmi (2.31.):

$$MSE_{reg} = \gamma \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 + (1-\gamma) \frac{1}{n} \sum_{j=1}^{n} w_j^2, \tag{2.31.}$$

kur γ – regularizācijas parametrs;

N – novērojumu skaits;

y – novērotā laika rinda;

$\hat{y}$ – prognozētā laika rinda;

n – svaru un konstanšu skaits;

w – svaru un konstanšu vektors.

Beijesa regularizācija (2.31.) minimizē kļūdu un svaru kvadrātu lineāro kombināciju, kas izraisa to, ka neironu tīkla modelim ir mazāki svari un konstantes, tas ļauj iegūt gludākas prognozes un samazina pārliekas pielāgotības (*overfitting*) iespēju prognozējamiem datiem [23].

Lai varētu izmantot Levenberga-Markvardta kļūdas atgriezeniskās izplatīšanās algoritmam ar Beijesa regularizāciju, neironu tīklā izmantotajām pārraides funkcijām ir jābūt diferencējamām. Savukārt konkrētiem datiem piemērots regularizācijas parametrs tiek atrast neironu tīkla modeļa apmācības laikā.

Lai atjaunotu svaru un konstantu vektoru katrā apmācības iterācijā, tiek izmantota Jakobiāna matrica un tuvināta Hesiana matrica [44]. Jakobiāna matricu, kas satur kļūdu e atvasinājumus pēc neironu tīkla modeļa svariem un konstantēm w, iegūst pēc vienādojuma (2.32.):

$$J = \begin{bmatrix} \frac{\partial e_1(w)}{\partial w_1} & \frac{\partial e_1(w)}{\partial w_2} & \cdots & \frac{\partial e_1(w)}{\partial w_n} \\ \frac{\partial e_2(w)}{\partial w_1} & \frac{\partial e_2(w)}{\partial w_2} & \cdots & \frac{\partial e_2(w)}{\partial w_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_N(w)}{\partial w_1} & \frac{\partial e_N(w)}{\partial w_2} & \cdots & \frac{\partial e_N(w)}{\partial w_n} \end{bmatrix}. \tag{2.32.}$$

Tuvinātu Hesiana matricu H iegūst pēc izteiksmes (2.33.):

$$H = J^T J\ . \tag{2.33.}$$

Svaru un konstanšu vektors w^{l+1} $l+1$ iterācijā var tikt atjaunots, izmantojot izteiksmi (2.34.):

$$w^{l+1} = w^l - [H + \mu I]^{-1} J^T e\,, \tag{2.34.}$$

kur w^l – svaru un konstanšu vektors *l-tajā* iterācijā;

H – Hesiana matrica;

μ – Levenberga faktors;

I – identitātes matrica;

J^T – transponētā Jakobiāna matrica;

e – kļūdu vektors, $e = y - \hat{y}$;

$J^T e$ – gradients.

Kad Levenberga faktors μ ir neliels, tad Levenberga-Markvardta kļūdas atgriezeniskās izplatīšanās algoritms ar Beijesa regularizāciju izmanto tuvināto Hesiana matricu. Ja μ ir nulle, tad tā ir Ņūtona metode, kas izmanto tuvināto Hesiana matricu. Savukārt ja μ ir liels, tad gradienta virzienā tiek veikti nelieli soļi, un algoritms paliek par gradienta nolaišanās kļūdas atgriezeniskās izplatīšanās algoritmu ar mazu soli [52].

2.5.2. Rekurento slāņu neironu tīkli

Rekurento slāņu neironu tīkli (*LRNN*) ir viens no dinamisko, rekurento neironu tīklu veidiem, kas ir veidoti uz Elmana neironu tīkla bāzes [19], [71]. *LRNN* vispārina Elmana rekurento neironu tīklu, ļaujot izmantot patvaļīgu slēpto slāņu skaitu, un patvaļīgas aktivizācijas funkcijas katrā slānī [71]. *LRNN* tiek veidots uz daudzslāņu

perceptrona bāzes, pievienojot atgriezeniskās saites [18]. Šīs saites iziet no slēpto slāņu neironiem, un nonāk konteksta slānī, kas ir rekurento slāņu neironu tīkla neironu tīkla papildus slānis. Konteksta slānis uzglabā slēpto slāņu vērtības ar laika aizturi, tādējādi sniedzot vērtīgu informāciju par iepriekšējo neironu tīkla ieejas vektoru [18]. Konteksta slāņa neironu skaits ir vienāds ar slēptā slāņa neironu skaitu. Rekurento slāņu neironu tīklam ar vienu slēpto slāni *k-to* ieejas vektoru izsaka ar izteiksmi (2.35.):

$$x(k) = \left[x_1(k), \ldots, x_N(k), v_1(k-1), v_2(k-1), \ldots, v_K(k-1) \right], \tag{2.35.}$$

kur N – ieejas atribūtu skaits;

K – konteksta slāņa neironu skaits;

k – neironu tīkla ieejas vektora indekss.

Tādējādi neironu tīkls aprēķina izejas signālu, balstoties ne tikai uz pašreizējo ieejas signālu, bet arī uz iepriekšējo ieejas signālu, un šī īpašība nosaka rekurento slāņu neironu tīkla galveno īpašību – secības iegaumēšanu [19]. Tas ļauj ņemt vērā novērojamo procesu vēsturi un iegūt labāku prognozēšanas precizitāti.

2.6. Izstrādājamās sistēmas precizitātes salīdzināšanai izmantotās prognozēšanas metodes

Lai varētu salīdzināt izstrādājamās prognozēšanas sistēmas precizitāti, darbā tiek izmantotas trīs prognozēšanas metodes: vienkāršais slīdošais vidējais (*Simple Moving Average, SMA*), diskrētā laika, nepārtrauktu stāvokļu kopas *m*-tās kārtas Markova ķēdes un automātiski regresējošs integrēts slīdošais vidējais (*Autoregressive Integrated Moving Average, ARIMA*).

2.6.1. Vienkāršais slīdošais vidējais

Vienkāršais slīdošais vidējais (*SMA*) ir funkcija, kuras vērtība katrā punktā ir vienāda ar vidējo aritmētisko vērtību noteiktā periodā [17]. Laika rindas y prognozi laika solim $t+1$ apraksta vienādojums (2.36.):

$$\hat{y}(t+1) = \frac{1}{m}\left(\sum_{i=t-m+1}^{t} y(i) \right) + \varepsilon(t+1), \tag{2.36.}$$

kur $\hat{y}(t+1)$ – laika rindas y prognozētā vērtība laika solī $t+1$;

$y(i)$ – laika rindas y novērotā vērtība laika solī i, $i = t - m + 1, \ldots, t$;

m – *SMA* periods;

$\varepsilon(t+1)$ – baltais troksnis laika solī $t+1$.

Ja no laika rindas tiek rekonstruēta fāžu telpa (skat. (2.23.)) ar noteiktu dimensiju (loga garumu), tad prognoze tiek aprēķināta kā katra fāžu telpas aiztures vektora (laika rindas fragmenta) vidējā vērtība.

2.6.2. Nepārtrauktu stāvokļu kopas Markova ķēdes

Izmantojot nepārtrauktu stāvokļu kopas Markova ķēdi, kuras kārta m ir Markova modeļa parametrs, kas apzīmē atmiņu, pašreizējais Markova procesa stāvoklis, X_t, var tikt aizvietots ar laika rindas pēdējiem m novērojumiem [86]. Laika rindas vērtības $y(t+1)$ atkarībā no pašreizējās un pagātnes novērojumu vērtībām $y(t), y(t-1), y(t-2), \ldots, y(t-m+1)$ var tikt aprakstīta, izmantojot nosacījumu varbūtību blīvuma funkciju [41], un varbūtība novērot vērtību v nākamajā periodā tiek aproksimēta ar vienādojumu (2.37.):

$$\Pr\{y(t+1) = v | y(t), y(t-1), \ldots, y(t-m+1)\}. \qquad (2.37.)$$

Izveidotajā Markova ķēdē katrs stāvoklis ir vienāds ar reālu skaitli. Tā kā tiek izmantota *m*-tās kārtas Markova ķēde, pēdējie *m* stāvokļi jeb pēdējie *m* laika rindas novērojumi tiek izmantoti, lai prognozētu nākamo stāvokli. Šī Markova ķēdes m stāvokļu kombinācija izveido vektoru, kas formāli ir identisks stāvokļa vektoram rekonstruētajā fāžu telpā (2.28.).

Ģeometriski pašreizējais stāvokļa vektors S_t ir punkts *m*-dimensiju fāžu telpā. Kaimiņu punkti šajā fāžu telpā parāda līdzīgus stāvokļu vektorus. Var uzskatīt, ka līdzīgiem punktiem fāžu telpā piemīt līdzīgs varbūtību blīvuma sadalījums [41]. Mainīgais $\Phi_\varepsilon(S_t)$ attēlo apkārtni ap vektoru S_t, kuru ierobežo neliels diametrs ε. Mainīgais $|\Phi_\varepsilon(S_t)|$ parāda vektoru $S_k, k < t$, skaitu, kas atrodas šajā apkārtnē. Vektori S_k tiek iegūti no laika rindas y pagātnes novērojumiem. Šiem vektoriem S_k nākamās vērtības nākotnē tiek apzīmētas ar v_{k+1}. Tad optimālā prognoze, kas minimizē zaudējumu funkcijas *RMSE* vērtību, var tikt iegūta kā nosacījumu varbūtības pirmais moments – vidējā vērtība, ko iegūst pēc izteiksmes (2.38.):

$$E\left(v_{k+1}\right) \approx \frac{1}{\left|\Phi_{\varepsilon}\left(S_t\right)\right|} \sum_{k \in \Phi_{\varepsilon}(S_t)} v_{k+1} \cdot \qquad (2.38.)$$

Nosacījumu varbūtības (2.38.) pirmais moments ir pašreizējā stāvokļa S_t iespējamo nākotnes vērtību vidējā vērtība – laika rindas y prognozētā vērtība laika solī $t+1$.

2.6.3. Automātiski regresējošs integrēts slīdošais vidējais

Automātiski regresējošs integrēts slīdošais vidējais (*ARIMA*) ir laika rindu analīzes un prognozēšanas metode. *ARIMA* modelis ir automātiski regresējoša slīdošā vidējā (*Autoregressive Moving Average*, *ARMA*) modeļa paplašinājums, ko var izmantot arī nestacionāru laika rindu modelēšanai un prognozēšanai. Lai iegūtu stacionāru laika rindu, *ARIMA* metode veic oriģinālās, nestacionārās laika rindas diferencēšanu, turklāt diferencēšana var tikt veikta ar dažādām kārtām.

Nestacionārās laika rindas y prognozi laika solim $t+1$, izmantojot $ARIMA(p,d,q)$ modeli [3] apraksta vienādojums (2.39.):

$$\Delta\hat{y}\left(t+1\right) = c + \sum_{i=1}^{p} a_i \Delta y\left(t-i+1\right) + \sum_{j=1}^{q} b_j \varepsilon\left(t-j+1\right) + \varepsilon\left(t+1\right), \qquad (2.39.)$$

kur $\Delta\hat{y}\left(t+1\right)$ – diferencētās laika rindas Δy ar integrācijas (diferenču operatora) kārtu d prognozētā vērtība laika solī $t+1$;

c – konstante;

p – autoregresīvā polinoma kārta;

a_i – autoregresijas i-*tais* koeficients; $i = 1,\ldots,p$;

$\Delta y\left(t-i+1\right)$ – diferencētās laika rindas Δy ar integrācijas kārtu d novērotā vērtība laika solī $t-i+1$;

q – slīdošā vidējā polinoma kārta;

b_j – slīdošā vidējā j-*tais* koeficients; $j = 1,\ldots,q$;

$\varepsilon\left(t-j+1\right)$ – atlikums (baltais troksnis) laika solī $t-j+1$.

Lai pārbaudītu, vai ieejas laika rinda ir stacionāra, tiek veikts vienības saknes tests – paplašinātais Dikija-Fullera tests [3]. Dikija-Fullera tests pārbauda nulles hipotēzi par vienības saknes eksistēšanu. Ja nulles hipotēze tiek noraidīta, tad laika rinda ir stacionāra, un to nav nepieciešams diferencēt. Ja tests nespēj noraidīt nulles hipotēzi par vienības saknes eksistēšanu, tad laika rindu ir nepieciešams diferencēt stacionāras laika rindas iegūšanai.

2.7. Kopsavilkums un secinājumi

Nodaļā dots ieskats par prognozēšanas sistēmas izstrādāšanā lietotajām datu pirmapstrādes un prognozēšanas metodēm. Pirms laika rindu prognozēšanas procesa tiek veikta datu pirmapstrāde, kurā notiek analizējamo datu kopu sagatavošana. Galvenie secinājumi par izstrādājamajā sistēmā izmantotajām metodēm ir šādi:

1. Laika rindas dekompozīcijai tiek piedāvāts izmantot variāciju apakšsignālu dekompozīcijas metodi – salīdzinoši jaunu, uz spektrālo analīzi balstītu signālu dekompozīcijas metodi, kas pieder pazīmju izvilkšanas metožu grupai, un, kura ļauj no skalārās laika rindas iegūt noteiktu skaitu apakšsignālu, kur katrs apakšsignāls atbilst noteiktām frekvencēm, un sniedz papildu informāciju par ieejas laika rindu.
2. Lai no skalārās laika rindas iegūtu ieejas datu kopu, var tikt izmantota fāžu telpas rekonstrukcija ar laika aiztures metodi.
3. No pazīmju atlasīšanas metodēm tiek piedāvāts izmantot pakāpenisko regresijas analīze – plaši izmantotu metodi regresijas analīzē un laika rindu prognozēšanas uzdevumos, kas no atribūtu kopām atlasa tādas pazīmju jeb atribūtu apakškopas, kas ļauj sasniegt augstāku prognozēšanas precizitāti.
4. Pakāpeniskās regresijas analīzes galvenais trūkums ir tas, ka tā nav globāli optimāla metode, un iegūtais regresijas modelis var būt tikai lokāli optimāls.
5. Pazīmju izvilkšanas metode, ko tiek piedāvāts izmantot izstrādājamajā sistēmā, ir principiālo komponentu analīze – šī ir populāra metode laika rindu pirmapstrādē, kas ļauj iegūt lineāri nekorelējošas atribūtu kopas, turklāt var tikt reducēta sākotnējo atribūtu kopu dimensiju skaits.

6. No laika rindu prognozēšanas metodēm tiek piedāvāts izmantot rekurento slāņu neironu tīklus (*LRNN*), kas ir Elmana rekurento neironu tīklu vispārinājums.
7. Kā *LRNN* apmācības algoritmu tiek piedāvāts izmantot Levenberga-Markvardta kļūdas atgriezeniskās izplatīšanās algoritmu ar Beijesa regularizāciju, kuram piemīt ātrāka konverģence, un kurš var atrast labāku lokālo minimumu, salīdzinot ar gradienta nolaišanās metodi.
8. Rekurento slāņu neironu tīkliem ir vairāki mīnusi, piemēram, ilgs apmācības laiks, resursu prasīgs un necaurredzams apmācības process, sarežģīta piemērotas arhitektūras un parametru vērtību izvēle.

Tomēr *LRNN*, pateicoties savām pozitīvajām īpašībām – spējai modelēt un ar labu precizitāti prognozēt tieši nelineāras, nestacionāras un trokšņainas laika rindas, kā arī spējai iegaumēt elementu secību, var uzrādīt augstāku prognozēšanas precizitāti *NDVI* indeksa laika rindu prognozēšanā, salīdzinot ar lineārām metodēm vai vienvirziena neironu tīkliem.

3. APAKŠSIGNĀLA APROKSIMĀCIJAS PIEJAS IZSTRĀDĀŠANA

Variāciju apakšsignālu dekompozīcijas (*Variational Mode Decomposition, VMD*) metode [118] pieder pie pazīmju izvilkšanas metožu grupas, līdz ar to šo metodi var pielietot *NDVI* laika rindu prognozēšanas sistēmā kā datu pirmapstrādes metodi. Respektīvi, ieejas laika rinda y tiek sadalīta kāda noteikta skaita apakšsignālos: $u_1, u_2, \ldots, u_K$, kur K ir apakšsignālu skaits. Gan oriģinālajai laika rindai y, gan visiem aprēķinātajiem apakšsignāliem $u_1, u_2, \ldots, u_K$ tiek iegūtas rekonstruētās fāžu telpas. Visas šīs fāžu telpas var tikt izmantotas kā ieejas dati, un var tikt meklētas funkcionālās sakarības starp oriģinālās laika rindas y novērojuma vērtību laika solī $t+1$ un šīs laika rindas un apakšsignālu novērojumu vērtībām laika soļos no $t-m+1$ līdz t atbilstoši uzdevuma formālajai nostādnei (skat. izteiksmi (1.7.)).

Kā tika apskatīts 1.10. apakšnodaļā, daudzos laika rindu prognozēšanas pētījumos, kur izmanto uz spektrālo analīzi balstītās laika rindu dekompozīcijas metodes, tajā skaitā *VMD* metodi, dekompozīcija tiek pielietota divos veidos.

Pirmajā variantā [33], [95] tā tiek pielietota visiem laika rindas vēsturiskajiem novērojumiem, un tikai tad no oriģinālās laika rindas un iegūtajiem apakšsignāliem tiek iegūtas rekonstruētās fāžu telpas, izmantojot laika aiztures metodi. Apvienojot fāžu telpas, tiek iegūta sākotnējā ieejas datu kopa, kas sastāv no atribūtu kopas un prognozējamā parametra. Ieejas datu kopa tiek sadalīta apmācības, validācijas un testa datu kopās, tiek veikta datu pirmapstrāde un tiek iegūts prognozēšanas modelis.

Otrajā variantā [30], [34] oriģinālā laika rinda tiek sadalīta fragmentos, atbilstoši noteiktām loga garumam m, izmantojot fāžu telpas rekonstrukciju ar laika aiztures metodi, un tikai tad *VMD* metode tiek pielietota katram laika rindas fragmentam, iegūstot atribūtu skaitu, kas ir vienāds ar loga garuma (fāžu telpas dimensijas) un apakšsignālu skaita K reizinājumu. Un tālāk gan šī atribūtu kopa, gan prognozējamais parametrs tiek sadalīti apmācības, validācijas un testa datu kopās, tiek veikta datu pirmapstrāde un prognozēšanas modeļa iegūšana.

Lai pārbaudītu, vai oriģinālā *VMD* metode ir izmantojama normalizēta veģetācijas indeksa (*Normalized Difference Vegetation Index, NDVI*) laika rindu prognozēšanas uzdevumā, un kādu prognozēšanas precizitāti var sasniegt, izmantojot šo metodi, tika veikti divi eksperimenti, kas tiek aprakstīti 3.1. apakšnodaļā.

3.1. Oriģinālās *VMD* metodes izmantošana *NDVI* laika rindu prognozēšanai

Oriģinālās variāciju apakšsignālu dekompozīcijas (*VMD*) metodes precizitātes novērtēšanai *NDVI* laika rindu prognozēšanas uzdevumā ir izstrādāts eksperimentu plāns, kas ir aprakstīts 3.1.1. punktā.

3.1.1. Eksperimentu plāns ar oriģinālo VMD metodi

Plānā ir iekļauti divi eksperimenti. Katra eksperimenta mērķis un apraksts *NDVI* laika rindu prognozēšanas uzdevumā ir dots 3.1. tabulā. Abos eksperimentos tiek izmantotas 100 *NDVI* laika rindas, kur katrai laika rindai prognozēšana tiek veikta atsevišķi. Laika rindas tiek izvēlētas, ģenerējot gadījuma skaitļus – pikseļu rindas un kolonnas numurus vidējās izšķirtspējas attēlveidošanas spektroradiometra (*Moderate Resolution Imaging Spectroradiometer, MODIS*) *NDVI* attēlos. Novērtēšana tiek veikta, aprēķinot trīs dažādu zaudējumu funkciju: saknes vidējās kvadrātiskās kļūdas (*Root Mean Square Error, RMSE*), virzienu simetrijas (*Directional Symmetry*, DS), un determinācijas koeficienta R^2 vērtības starp novēroto un prognozēto *NDVI* laika rindu.

3.1. tabula

Eksperimentu plāns ar oriģinālās *VMD* metodes izmantošanu

Nr.	Eksperimenta apraksts	Eksperimenta mērķis
1.	Iegūt apakšsignālus, oriģinālās *VMD* metodes ieejā, padodot visus *NDVI* vēsturiskos novērojumus. Veikt *NDVI* laika rindas prognozēšanu, izmantojot apakšsignālus un *NDVI* novērojumus.	Pārbaudīt *NDVI* laika rindu prognozēšanas precizitāti, pielietojot oriģinālo *VMD* metodi visiem *NDVI* laika rindas vēsturiskiem novērojumiem.
2.	Iegūt apakšsignālus, oriģinālās *VMD* metodes ieejā, padodot katru *NDVI* laika rindas fragmentu. Veikt *NDVI* laika rindas prognozēšanu, izmantojot apakšsignālus un *NDVI* novērojumus.	Pārbaudīt *NDVI* laika rindu prognozēšanas precizitāti, pielietojot oriģinālo *VMD* metodi katram *NDVI* laika rindas fragmentam.

Pirmajā eksperimentā tiek novērtēta prognozēšanas precizitāte, izmantojot oriģinālo *VMD* metodi visiem vēsturiskajiem *NDVI* laika rindas novērojumiem. Otrajā eksperimentā tiek pārbaudīta prognozēšanas precizitāte, izmantojot oriģinālo *VMD* metodi katram *NDVI* laika rindas fragmentam, kas iegūti no rekonstruētās fāžu telpas ar noteiktu dimensiju (loga garumu).

3.1.2. Oriģinālās VMD metodes izmantošana visiem vēsturiskajiem novērojumiem

Eksperimentā tiek izmantotas 100 *NDVI* laika rindas, kas tiek izvēlētas, ģenerējot gadījuma skaitļus, kas atbilst pikseļa rindas un kolonnas numuriem *MODIS NDVI* attēlos. Katra laika rinda satur 814 novērojumus. Eksperiments tiek atkārtots 100 reižu, katru reizi izmantojot vienu no šīm *NDVI* laika rindām.

Tā kā, izmantojot oriģinālo *VMD* metodi uz visiem vēsturiskajiem datiem, ir iespējams prognozēt tikai vienu jaunu laika rindas vērtību, kas atrodas vienu laika soli aiz vēsturiskajiem datiem, tad, lai testa datu kopa nesastāvētu tikai no viena datu ieraksta, no katras laika rindas secīgi 245 reizes pēc kārtas tiek atlasīti 70 % novērojumu (novērojumu skaits $N = 570$), pirmajā iterācijā atlasot 570 novērojumus no pirmā novērojuma līdz 570. novērojumam, otrajā iterācijā atlasot 570 novērojumus no otrā novērojuma līdz 571. novērojumam, līdz pēdējā iterācijā tiek atlasīti 570 novērojumi no 245. novērojuma līdz 814. novērojumam, kā tas parādīts 3.1. attēlā.

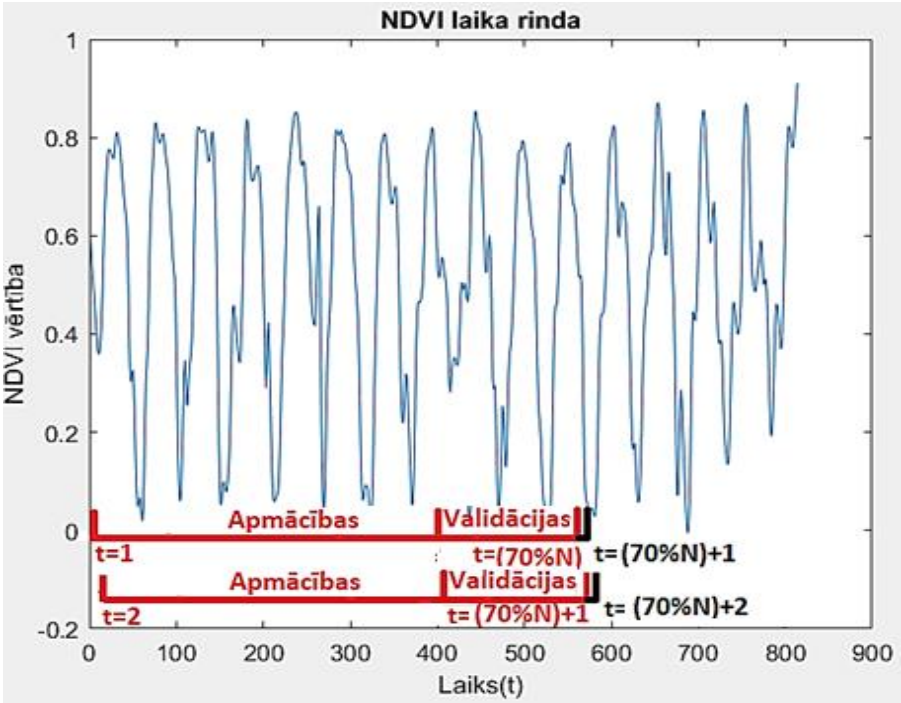
Prognozējot pirmo *NDVI* laika rindu no 100 laika rindām, pirmajā iterācijā atlasītie 70 % novērojumu tiek padoti ieejā oriģinālajai *VMD* metodei. Oriģinālajai *VMD* metodei [119] ieejā tiek padoti septiņi parametri, kas nosaka gan apakšsignālu skaitu, gan to vērtības:

y – laika rinda, kurai tiek veikta dekompozīcija;
α – apakšsignālu summas precizitāti regulējošs parametrs;
τ – laika solis;
K – apakšsignālu skaits;
ε – konverģences kritērijs;
DC – parametrs, kuram divas iespējamās vērtības:

- 1, ja pirmajam apakšsignālam atbilstošas aprēķinātās centrālās frekvences ω_1 var būt tikai nulles;
- 0, ja pirmajam apakšsignālam atbilstošās aprēķinātās centrālās frekvences ω_1 var pieņemt dažādas vērtības;

init – centrālo frekvenču inicializācijas veids, kuram ir trīs iespējamās vērtības:

- 0, ja frekvences tiek inicializētas ar nullēm;
- 1, ja frekvences tiek inicializētas ar vienmērīga sadalījuma skaitļiem, *k-tām* apakšsignālam sākotnēji katra centrālā frekvence no ω_k ir vienāda ar $0.5K(k-1)$;
- 2, ja frekvences tiek inicializētas ar gadījuma skaitļiem.



3.1. att. Novērojumu atlasīšana.

Laika soli τ oriģinālās *VMD* metodes autori rekomendē ņemt kā nulles vērtību, vai kā nelielu skaitli, nepārsniedzot vērtību 0,1 [118]. Šajā eksperimentā darba autors laika solim τ izmanto vērtību 0,000001. Parametram DC tiek uzstādīta vērtība 1, tādējādi centrālās frekvences tiek inicializētas ar nullēm, un konverģences kritērija ε vērtība arī tiek uzstādīta kā nulles vērtība. Piemērots apakšsignālu summas precizitāti regulējošs parametrs α un apakšsignālu skaits K katrai laika rindai tiek atrasti eksperimenta laikā. Parametrs α tiek meklēts intervālā [1; 100], lielās α vērtības noved pie tā, ka apakšsignālu summa sāk atšķirties no oriģinālās *NDVI* laika rindas. Apakšsignālu skaits K tiek meklēts intervālā [1; 6], kur maksimālais apakšsignālu skaits tiek noteikts, aprēķinot veselu skaitli no $\log(N)$.

NDVI laika rindai un katram apakšsignālam tiek veikta datu pirmapstrāde, rekonstruējot fāžu telpu ar laika aiztures metodi, kur dimensija $m = 50$ un laika

aizture $\tau = 1$. Turklāt fāžu telpa tiek rekonstruēta gan oriģinālajai *NDVI* laika rindai, gan katram apakšsignālam, iegūstot $K+1$ rekonstruētās fāžu telpas. Apvienojot rekonstruētās fāžu telpas, tiek iegūta sākotnējā ieejas datu kopa, kura tiek sadalīta divās daļās pēc datu ierakstiem: 80 % no ieejas datu kopas ierakstu skaitā veido apmācības datu kopu un 20 % validācijas datu kopu. Tad katra no šīm divām datu kopā tiek sadalīta divās daļās: atribūtu (pazīmju) kopā, kuru veido visi datu kopas (apmācības, validācijas vai testa) ieraksti no pirmā datu ieraksta līdz pirmspēdējam, un prognozējamajā parametrā, par kuru tiek izmantots no *NDVI* laika rindas rekonstruētās fāžu telpas pēdējais (piecdesmitais) atribūts, kura elementus veido datu ieraksti no otrā līdz pēdējam datu ierakstam.

Tad, izmantojot pakāpenisko regresijas analīzi, no sākotnējiem atribūtiem apmācības datu kopā tiek atlasīti informatīvākie atribūti, iegūstot šo atribūtu indeksus. Tad atribūti ar šādiem indeksiem tiek atlasīti arī no validācijas atribūtu kopas. Reducētās apmācības atribūtu kopa tiek padota ieejā principiālo komponentu analīzei, iegūstot tādu īpašvektoru kopu (matricu), kuru sareizinot ar apmācības atribūtu kopu, iegūst lineāri nekorelējošu apmācības datu kopas atribūtu kopu. Šī īpašvektoru kopa tiek izmantota, lai arī validācijas atribūtu kopā iegūtu lineāri nekorelējošu atribūtu kopu.

Savukārt par prognozēšanas metodi tiek izmantots rekurento slāņu neironu tīkls (*Layer Recurrent Neural Network, LRNN*), kas ieejā saņem apmācības un validācijas datu kopas pēc datu pirmapstrādes veikšanas.

- *LRNN* modeļu svaru un konstanšu vērtības tiek inicializētas ar gadījuma skaitļiem intervālā [–0,1; 0,1].
- Apmācība tiek veikta ar Levenberga-Markvardta kļūdas atgriezeniskās izplatības algoritmu ar Beijesa regularizāciju.
- Slēpto slāņu skaits ir viens slānis visiem neironu tīkliem.
- Hiperboliskā tangensa funkcija un lineārā funkcija tiek izmantotas kā aktivizācijas funkcijas atbilstoši slēptā un izejas slāņa neironiem.
- Epohu skaits, kas tiek izmantotas *LRNN* apmācībai, ir 10000.

Katram *LRNN* modelim piemērots slēptā slāņa neironu skaits, kas dod mazāko *RMSE* vērtību uz validācijas datu kopas, tiek atrasts, izmantojot šķērsvalidāciju. Neironu skaits slēptajā slānī tiek meklēts intervālā [1; 50]. Lai uzlabotu neironu tīkla vispārināšanas spējas, tiek izmantota agrīnās apstāšanās metode. Kad kļūda uz

validācijas datu kopas pieaug 100 epohas pēc kārtas, apmācība tiek apturēta, un tiek paņemtas tās svaru un konstanšu vērtības, kas tika iegūtas pie minimālās kļūdas vērtības, kas tika sasniegta 101 epohu atpakaļ.

Kad prognozēšanas modelis ir iegūts, tas tiek izmantots, lai prognozētu vienu nākamo *NDVI* laika rindas vērtību, kas atrodas aiz šiem 70 % novērojumu, kas pirmajā iterācijā ir 571. vērtība. Tad tiek paņemti nākamie 70 % novērojumu, no 2. novērojuma līdz 571. novērojumam un visa datu pirmapstrādes un prognozēšanas modeļa apmācības procedūra tiek atkārtota, rezultātā prognozējot 572. vērtību. Kad ir izietas visas iterācijas, katrā no tām atlasot 70 % novērojumu, un, prognozējot vienu *NDVI* laika rindas vērtību uz priekšu, tiek iegūtas prognozes laika rindas vērtībām no 571. vērtības līdz 814. vērtībai, rezultātā iegūstot testa datu kopu, kas sastāda 30 % no visiem *NDVI* laika rindas novērojumiem. Tālāk tiek paņemta nākamā *NDVI* laika rinda, un eksperiments tiek atkārtots.

3.1.3. Oriģinālās VMD metodes izmantošana katram laika rindas fragmentam

Eksperimentā tiek izmantotas 100 *NDVI* laika rindas, kur katra satur 814 novērojumus. Katra laika rinda tiek izvēlēta, ģenerējot gadījuma skaitļus, kas ir pikseļa rindas un kolonnas numuri *MODIS NDVI* attēlos.

Prognozējot pirmo *NDVI* laika rindu no 100 laika rindām, tiek veikta fāžu telpas rekonstrukcija ar laika aiztures metodi, kur dimensija $m = 50$, bet laika aizture $\tau = 1$. Rekonstruētā fāžu telpa tiek sadalīta divās daļās: atribūtu kopā, kuru veido fāžu telpas ieraksti no pirmā līdz pirmspēdējam datu ierakstam, un prognozējamajā parametrā, kuru veido fāžu telpas pēdējā kolonna no otrā līdz pēdējam datu ierakstam. Katrs atribūtu kopas ieraksts, kas atbilst *NDVI* laika rindas fragmentam ar loga garumu 50, tiek padots ieejā oriģinālajai *VMD* metodei.

Rezultātā sākotnējā atribūtu kopa, kas satur 50 atribūtus, paplašinās līdz $50(K+1)$ atribūtiem, kur K ir apakšsignālu skaits, ko izmanto *VMD* metode. Par laika soļa τ vērtību šajā eksperimentā darba autors izmanto vērtību 0,000001. Parametram DC tiek uzstādīta vērtība 1, tādējādi centrālās frekvences tiek inicializētas ar nullēm, un konverģences kritērija ε vērtība arī tiek uzstādīta kā nulles vērtība. Piemērots apakšsignālu summas precizitāti regulējošs parametrs α un apakšsignālu skaits K visiem laika rindas fragmentiem vienlaikus tiek atrasts eksperimenta laikā.

Parametrs α tiek meklēts intervālā [1; 100], bet apakšsignālu skaits K tiek meklēts intervālā [1; 3], kur maksimālais apakšsignālu skaits tiek noteikts, aprēķinot veselu skaitli no $\log(m)$.

Paplašinātā atribūtu kopa un prognozējamais parametrs tiek sadalīti trīs daļās pēc datu ierakstiem: 70 % no datu ierakstu skaita veido apmācības datu kopu, 15 % validācijas datu kopu un 15 % testa datu kopu.

Izmantojot pakāpenisko regresijas analīzi, no sākotnējiem atribūtiem apmācības atribūtu kopā tiek atlasīti informatīvākie atribūti, iegūstot šo atribūtu indeksus, un atribūti ar atbilstošiem indeksiem tiek atlasīti arī no validācijas un testa atribūtu kopām. Reducētā apmācības atribūtu kopa tiek padota ieejā principiālo komponentu analīzei, iegūstot tādu īpašvektoru kopu, kas ļauj iegūt lineāri nekorelējošu apmācības atribūtu kopu. Šī īpašvektoru kopa tiek izmantota, lai arī no validācijas un testa atribūtu kopām iegūtu lineāri nekorelējošas atribūtu kopas. Par prognozēšanas metodi tiek izmantots rekurento slāņu neironu tīkls (*LRNN*).

- *LRNN* modeļu svaru un konstanšu vērtības tiek inicializētas ar gadījuma skaitļiem intervālā [–0,1; 0,1].
- Apmācībai tiek izmantots Levenberga-Markvardta kļūdas atgriezeniskās izplatības algoritms ar Beijesa regularizāciju.
- Izmantotajā neironu tīkla arhitektūrā slēpto slāņu skaits ir viens slānis.
- Slēptā slāņa neroniem kā aktivizācijas funkcija tiek izmantota hiperboliskā tangensa funkcija, bet izejas slānim tiek izmantota lineārā funkcija.
- Apmācība apstājas, ja tiek sasniegts epohu skaits 10000.

Izmantojot šķērsvalidāciju, katram *LRNN* modelim tiek atrasts piemērots slēptā slāņa neironu skaits, kas ļauj iegūt mazāko *RMSE* vērtību validācijas datu kopā. Piemērots neironu skaits slēptajā slānī tiek meklēts intervālā [1; 50]. Tiek izmantota agrīnās apstāšanās metode, ja kļūda uz validācijas datu kopas pieaug 100 epohas pēc kārtas, tad kā piemērotākais tiek paņemts tas *LRNN* modelis, kas tika iegūts 101 epohu atpakaļ.

Kad prognozēšanas modelis ir iegūts, tas tiek izmantots, lai prognozētu testa datu kopas prognozējamo parametru, kuru veido *NDVI* laika rindas vērtības. Kad prognozes tiek iegūtas, tad tālāk tiek paņemta nākamā *NDVI* laika rinda, un eksperiments tiek atkārtots.

3.1.4. Rezultātu analīze eksperimentiem ar oriģinālo VMD metodi

Kā prognožu zaudējumu funkcijas tika izmantotas saknes vidējā kvadrātiskā kļūda (*RMSE*), virzienu simetrija (DS) un determinācijas koeficients R^2. Pirmajā eksperimentā iegūtas zaudējumu funkciju vidējās vērtības no visām 100 laika rindām uz katras datu kopas dotas zemāk redzamajā 3.2. tabulā.

3.2. tabula

Zaudējumu funkciju vērtības eksperimentos ar oriģinālo *VMD* metodi

Datu kopa	**Pirmais eksperiments**			**Otrais eksperiments**		
	RMSE	**DS**	**R^2**	***RMSE***	**DS**	**R^2**
Apmācības datu kopa	0,000027	99,94%	1,00	0,0013	97,53%	0,99
Validācijas datu kopa	0,000040	99,99%	1,00	0,0016	97,45%	0,99
Testa datu kopa	0,046000	85,59%	0,96	0,0011	97,69%	0,99

Pirmajā eksperimentā iegūtās prognožu zaudējumu funkciju vērtības, kas parādītas 3.2. tabulā, uzrāda augstu prognozēšanas precizitāti uz apmācības un validācijas datu kopām. Šāda prognozēšanas precizitāte būtu pietiekami augsta, lai tādu prognozēšanas modeli varētu izmantot *NDVI* laika rindu prognozēšanai. Tomēr, kā redzams, uz testa datu kopām visu trīs zaudējumu funkciju vidējās vērtības uzrāda neapmierinošu prognozēšanas precizitāti. Līdz ar to šāds prognozēšanas modelis nebūtu piemērots praktiskai *NDVI* laika rindu prognozēšanai un reālo uzdevumu veikšanai.

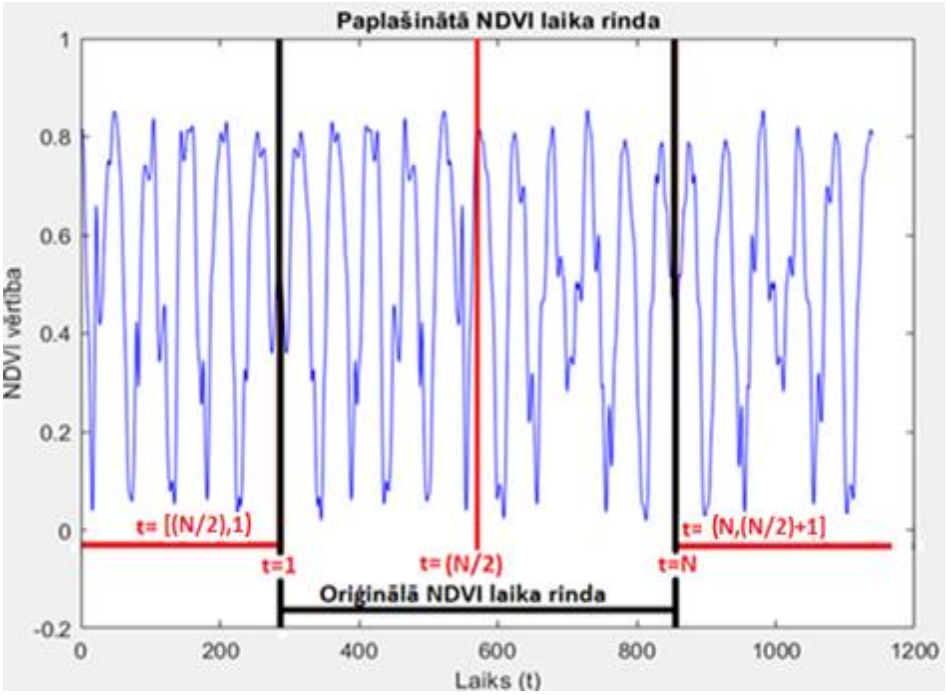
Tas ir saistīts ar to, ka variāciju apakšsignālu dekompozīcijas metode, aprēķinot apakšsignālu vērtības katrā laika solī t, izmanto visas ieejas signāla vērtības. Respektīvi, katra apakšsignāla vērtība katrā laika solī t satur informāciju par laika rindas vērtībām gan iepriekšējos, gan nākamajos laika soļos. Bet pēdējam ieejas laika rindas novērojumam laika solī $t = N$ nākamās vērtības nav pieejamas. Tā vietā oriģinālās *VMD* metodes autori izmanto signālu "spoguļošanu", paplašinot laika rindu ar vērtībām no vēsturiskajiem novērojumiem [118].

Otrajā eksperimentā aprēķinātas prognožu zaudējumu funkciju vidējās vērtības. Kā var redzēt 3.2. tabulā, tās uzrāda stabilu prognozēšanas precizitāti uz visām trim datu kopām atšķirībā no pirmajā eksperimentā iegūtās precizitātes, kur stabila un augsta prognozēšanas precizitāte tiek sasniegta tikai uz apmācības un validācijas datu kopām,

bet uz testa datu kopām precizitāte strauji samazinās. Salīdzinot abu eksperimentu prognozēšanas precizitāti uz apmācības un validācijas datu kopām, 3.2. tabulā ir redzams, ka augstāka precizitāte tiek sasniegta, izmantojot oriģinālo *VMD* metodi uz visiem vēsturiskajiem novērojumiem. Tas ir tādēļ, ka iegūtie apakšsignāli, izmantojot *VMD* metodi visiem vēsturiskajiem novērojumiem, satur vairāk informācijas par ieejas laika rindu.

3.1.5. Oriģinālās VMD metodes problēmas NDVI laika rindu prognozēšanā

Oriģinālajā *VMD* metodē izmantotais "spoguļošanu" algoritms ir šāds. Ja ieejas signāla garums ir N, tad tiek aprēķināts signāla viduspunkts: vesels skaitlis no $N/2$. No ieejas signāla tiek paņemti visi elementi no pirmā elementa līdz $N/2$ elementam, un apgrieztā secībā tiek pielikti priekšā ieejas signālam. Tālāk tiek paņemti visi elementi no $N/2+1$ elementa līdz N elementam un apgrieztā secībā tiek pielikti ieejas signāla beigās. Līdz ar to variāciju apakšsignālu dekompozīcijas metode strādā ar signālu, kura garums ir $2N$. Zemāk redzamajā attēlā (skat. 3.2. att.) ir parādīta ieejas signāla (dotajā gadījumā – *NDVI* laika rindas ar garumu $N = 814$) "spoguļošana", ko izmanto oriģinālajā variāciju apakšsignālu dekompozīcijas metodē.



3.2. att. Ar "spoguļošanas" palīdzību paplašinātā *NDVI* laika rinda.

Tomēr, skatoties no laika rindu prognozēšanas skatupunkta, šāda "spoguļošanas" pieeja rada noteiktas problēmas – paplašinātās laika rindas daļas arī

dod savu ieguldījumu apakšsignālu skaitļošanas procesā, un līdz ar to arī ietekmē, kādas būs apakšsignālu vērtības laika soļos no $t = 1$ līdz $t = N$. Tas noved pie problēmas, ka, izstrādājot prognozēšanas modeļus, tie spēj ar augstu precizitāti prognozēt tikai vēsturiskos laika rindas novērojumus. Ja ir nepieciešams prognozēt laika rindas novērojumus aiz laika soļa $t = N$, piemēram, $t = N + 1$ novērojuma vērtību, tad prognozēšanas precizitāte strauji krīt. Tas ir saistīts ar to, ka apakšsignāla vērtība laika solī $t = N$ satur informāciju par nākamo ieejas laika rindas elementu, kas, kā redzams 3.2. attēlā, ir paplašinātās laika rindas elements laika solī $t = N + 1$. Līdz ar to prognozētā vērtība, ar kādu noteiktu kļūdas lielumu, būs līdzīga ar paplašinātā signāla novērojuma vērtību atbilstošajā laika solī, un nevis ar īsto laika rindas novērojuma vērtību laika solī $t = N + 1$. Ar to ir izskaidrojama precizitātes strauja krišana, prognozējot vērtības aiz laika soļa $t = N$. Tādēļ ir nepieciešama *VMD* metodes modifikācija, kurā netiktu izmantota augstāk aprakstīta "spoguļošanas" pieeja, kas nodrošinātu to, ka šo metodi, lietojot to kopā ar kādu noteiktu pieeju, varētu izmantot jaunu *NDVI* laika rindu vērtību prognozēšanai.

3.2. *VMD* metodes modifikācijas izstrādāšana laika rindu prognozēšanai

Lai atrisinātu "spoguļošanas" izraisīto problēmu, viens no variantiem ir neizmantot laika rindu paplašināšanu, bet tā vietā divreiz samazināt oriģinālā signāla novērojumu skaitu, ko ieejā saņem *VMD* metode. Ja visa oriģinālā laika rinda (kas tagad tiek izmantota kā paplašinātā laika rinda) ir ar novērojumu skaitu $2N$, tad laika rindas vidusdaļa satur N novērojumus, un arī iegūtie apakšsignāli atbilst šiem N novērojumiem. Tagad paplašinājumu vietā atradīsies oriģinālās laika rindas reālie novērojumi, un apakšsignālu vērtības visos pieejamos laika soļos saturēs papildu informāciju par nākamajām laika rindas vērtībām, un tas ļaus prognozēt ar augstu precizitāti vienu laika rindas vērtību arī aiz laika soļa $t = N$.

Ja oriģinālās (ieejas) laika rindas novērojumu skaits $T = 570$ (tiek paņemti 70 % no pieejamiem 814 novērojumiem), tad viduspunkts, šajā gadījumā, atradīsies 285. novērojumā, ko iegūst aprēķinot formulu (3.1.):

$$B = T / 2 . \qquad (3.1.)$$

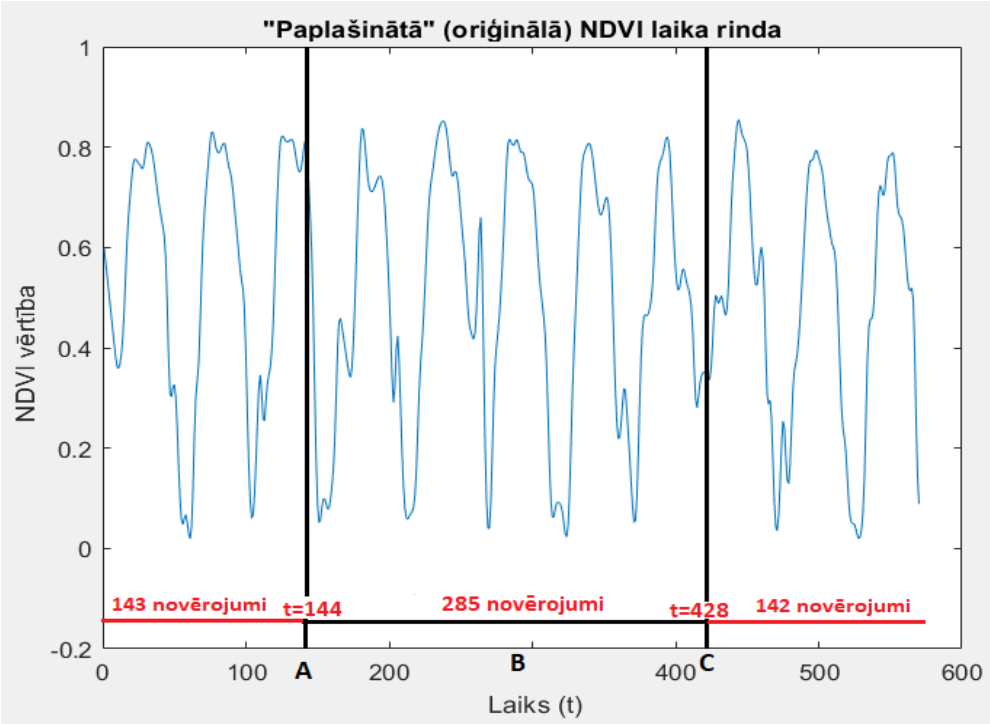
Attiecīgi paplašinājums laika rindas sākumā ietvers novērojumus no pirmā novērojuma līdz punktam A, ko iegūst pēc formulas (3.2.):

$$A = \frac{T}{4} + 1. \tag{3.2.}$$

Šajā apskatītajā piemērā paplašinājums laika rindas sākumā būs līdz 143. novērojumam, par jaunas "oriģinālās" *NDVI* laika rindas sākuma punktu uzskatot punktu A. Savukārt paplašinājums laika rindas beigās atbilst oriģinālās laika rindas novērojumiem no 429. novērojuma līdz 570. novērojumam un jaunas "oriģinālās" *NDVI* laika rindas beigu punktu C iegūst, aprēķinot formulu (3.3.):

$$C = 3\frac{T}{4} + 1. \tag{3.3.}$$

Novērojumu skaits jaunajai "oriģinālajai" *NDVI* laika rindai šajā gadījumā būs 285.



3.3. att. *NDVI* laika rindas sadalīšana "paplašinātajā" un "oriģinālajā" daļā.

Ar formulām (3.1.) ÷ (3.3.) aprēķināta "oriģinālās" un "paplašinātās" laika rindas proporcionalitāte atbilst oriģinālās un paplašinātās laika rindas proporcionalitātei, ko iegūst, izmantojot oriģinālo *VMD* metodi, respektīvi, oriģinālās laika rindas garums ir puse no paplašinātās laika rindas garuma, turklāt paplašinātas vērtības atrodas abos oriģinālās laika rindas galos.

Tomēr šāda modifikācija vēl nav izmantojama *NDVI* laika rindu jaunu vērtību prognozēšanai, jo iegūtās apakšsignālu vērtības ir pieejamas tikai oriģinālās ieejas laika rindas vidusdaļai, kā tas ir redzams 3.3. attēlā. Attiecīgi arī laika rindas vērtība, ko var prognozēt vienu laika soli uz priekšu, atradīsies aiz šīs vidusdaļas novērojumiem, laika solī $t = N + 1$, kamēr *NDVI* prognozēšanas uzdevumā, ir nepieciešams prognozēt $t = T + 1$ vērtību.

Tomēr šī *VMD* metodes modifikācija izejā dod apakšsignālus, kas var tikt izmantoti par pamatu pieejas izstrādāšanai, kas ļaus lietot no *VMD* metodes iegūtos apakšsignālus *NDVI* laika rindu prognozēšanas uzdevumā. Tādēļ ir svarīgi, lai ar autora modificēto *VMD* metodi varētu tikt sasniegta augsta *NDVI* laika rindu prognozēšanas precizitāte, tajā skaitā vienam laika solim aiz "oriģinālās" laika rindas vēsturiskiem novērojumiem.

Modificētas variāciju apakšsignālu dekompozīcijas (*VMD*) metodes precizitātes novērtēšanai *NDVI* laika rindu prognozēšanas uzdevumā ir izstrādāts eksperimentu plāns, kas ir aprakstīts 3.3. apakšnodaļā.

3.3. Eksperimentu plāns ar modificēto *VMD* metodi

Plānā ir iekļauti divi eksperimenti, kas ir apzīmēti kā trešais un ceturtais eksperiments. Katra eksperimenta mērķis un apraksts *NDVI* laika rindu prognozēšanas uzdevumā ir dots 3.3. tabulā. Trešajā eksperimentā tiek novērtēta prognozēšanas precizitāte, izmantojot modificēto *VMD* metodi visiem vēsturiskiem *NDVI* laika rindas novērojumiem. Ceturtajā eksperimentā tiek pārbaudīta prognozēšanas precizitāte, izmantojot no modificētās *VMD* metodes iegūto apakšsignālu *NDVI* laika rindu prognozēšanas uzdevumā un mainot šī apakšsignāla ciparu skaitu aiz komata.

Abos eksperimentos tiek izmantotas 100 *NDVI* laika rindas, kas arī tika izmantotas iepriekšējos divos eksperimentos, kas aprakstīti 3.1. apakšnodaļā, lai objektīvi novērtētu sasniegto *NDVI* laika rindu prognozēšanas precizitāti, izmantojot modificēto *VMD* metodi, un salīdzinātu to ar precizitāti, ko var sasniegt, izmantojot oriģinālo *VMD* metodi, tās ieejā padodot gan visus *NDVI* laika rindas vēsturiskos novērojumus, gan *NDVI* laika rindas fragmentus. Katra laika rinda sākotnēji satur 814 novērojumus. Laika rindas tiek izvēlētas, ģenerējot gadījuma skaitļus – pikseļu rindas un kolonnas numurus *MODIS NDVI* attēlos.

3.3. tabula

Eksperimentu plāns ar modificētas *VMD* metodes izmantošanu

Nr.	Eksperimenta apraksts	Eksperimenta mērķis
3.	Iegūt apakšsignālus, modificētās *VMD* metodes, padodot ieejā visus *NDVI* vēsturiskos novērojumus. Veikt *NDVI* laika rindas prognozēšanu, izmantojot iegūtos apakšsignālus un *NDVI* novērojumus.	Pārbaudīt prognozēšanas precizitāti, pielietojot modificēto *VMD* metodi visiem *NDVI* laika rindas vēsturiskiem novērojumiem.
4.	Iegūt apakšsignālus, modificētās *VMD* metodes, padodot ieejā visus *NDVI* vēsturiskos novērojumus. Visiem apakšsignāliem secīgi atmest daļu no ciparu skaita aiz komata. Veikt *NDVI* laika rindas prognozēšanu, izmantojot iegūtos dažādas precizitātes apakšsignālus un *NDVI* novērojumus.	Pārbaudīt, kā mainās *RMSE* vērtības atkarībā no apakšsignālu precizitātes.

Novērtēšana tiek veikta, aprēķinot trīs dažādu zaudējumu funkciju: saknes vidējās kvadrātiskās kļūdas *RMSE*, virzienu simetrijas *DS* un determinācijas koeficienta R^2 vērtības starp novēroto un prognozēto *NDVI* laika rindu.

3.3.1. NDVI laika rindu prognozēšana ar modificētās VMD metodes izmantošanu

Trešajā eksperimentā tiek paņemta pirmā laika rinda, un tiek veikta pirmā iterācija, kur no pirmās laika rindas tiek atlasīti pirmie 70 % novērojumu, no laika soļa $t = 1$ līdz $t = 570$, kas kopā ir 570 novērojumi $(T = 570)$. Padodot atlasītos 570 novērojumus ieejā modificētajai *VMD* metodei, apakšsignāli tiek iegūti laika soļiem no $t = 143$ līdz $t = 428$, un kopā katrs apakšsignāls satur 285 novērojumus $(N = 285)$.

Šajā eksperimentā darba autors laika solim τ izmanto nelielu vērtību 0,000001. Parametram DC tiek uzstādīta vērtība 1, tādējādi centrālās frekvences tiek inicializētas ar nullēm, un konverģences kritērija ε vērtība arī tiek uzstādīta nulle. Piemērots apakšsignālu summas precizitāti regulējošs parametrs α un apakšsignālu skaits K katrai laika rindai tiek atrasts eksperimenta laikā. Parametrs α tiek meklēts

intervālā [1; 100]. Apakšsignālu skaits K tiek meklēts intervālā [1; 5], kur maksimālais apakšsignālu skaits tiek noteikts, aprēķinot veselu skaitli no $\log(N)$.

Normalizēta veģetācijas indeksa "oriģinālajai" laika rindai (ar 285 novērojumiem) un katram apakšsignālam tiek veikta fāžu telpas rekonstrukcija ar laika aiztures metodi, kur fāžu telpas dimensija $m = 50$ un laika aizture $\tau = 1$. Rekonstruētās fāžu telpas tiek apvienotas, iegūstot sākotnējo ieejas datu kopu, kura tiek sadalīta divās daļas: atribūtu kopā un prognozējamajā parametrā. Atribūtu kopu veido visi ieejas datu kopas ieraksti no pirmā līdz pirmspēdējam. Prognozējamo parametru veido no *NDVI* laika rindas rekonstruētās fāžu telpas pēdējā (piecdesmitā) kolonna no otrā līdz pēdējam datu ierakstam. Gan atribūtu kopa, gan prognozējamais parametrs tiek sadalīti trīs daļās pēc datu ierakstiem: 70 % no ierakstu skaita veido apmācības datu kopu, 15 % validācijas un 15 % testa datu kopu.

Par prognozēšanas metodi tiek izmantots rekurento slāņu neironu tīkls (*LRNN*). Apmācība tiek veikta uz apmācības datu kopas, piemēroti neironu tīkla arhitektūras parametri tiek meklēti uz validācijas datu kopas.

- *LRNN* modeļu svaru un konstanšu vērtības tiek inicializētas ar gadījuma skaitļiem intervālā [–0,1; 0,1].
- Apmācībai tiek izmantots Levenberga-Markvardta kļūdas atgriezeniskās izplatības algoritms ar Beijesa regularizāciju.
- Izmantotajā *LRNN* arhitektūrā ir viens slēptais slānis.
- Hiperboliskā tangensa funkcija tiek izmantota kā aktivizācijas funkcija slēptā slāņa neironiem, bet lineārā funkcija – izejas slāņa neironiem.
- Apmācības iterāciju skaits ir 10000.

Izmantojot šķērsvalidāciju, katram *LRNN* modelim tiek atrasts piemērots slēptā slāņa neironu skaits, kas dod mazāko *RMSE* kļūdu uz validācijas datu kopas. Neironu skaits slēptajā slānī tiek meklēts intervālā [1; 50]. Tiek izmantota agrīnās apstāšanās metode: kļūdai validācijas datu kopā pieaugot 100 epohas pēc kārtas, kā piemērotākais tiek paņemts *LRNN* modelis kas iegūts 101 epohu atpakaļ.

Kad prognozēšanas modelis ir iegūts, tas tiek izmantots, lai prognozētu *NDVI* laika rindas vērtību laika solī $t = N + 1$. Kad prognoze tiek iegūta, eksperimenta pirmā iterācija pirmajai laika rindai tiek pabeigta, un tālāk tiek paņemti nākamie 70 % novērojumi no laika soļa $t = 2$ līdz laika solim $t = T + 1$, un tiek veikta eksperimenta otrā iterācija pirmajai laika rindai, prognozējot *NDVI* vērtību laika solī $t = N + 2$.

NDVI laika rindas novērojumi no $t = N + 1$ līdz $t = T$ veido testa datu kopu. Kad visi testa datu kopas prognozējamā parametra elementi tiek prognozēti, tad tiek paņemta nākamā *NDVI* laika rinda, un eksperiments tiek atkārtots.

3.3.2. NDVI laika rindu prognozēšana ar dažādu apakšsignāla precizitāti

Lietojumprogramma *MATLAB* savos aprēķinos izmanto dubulto precizitāti ar 16 cipariem aiz komata. Tuvināti aprēķinot apakšsignālus, var tikt iegūtas vērtības ar dažādu ciparu skaitu aiz komata, salīdzinot ar oriģinālo apakšsignālu. Respektīvi, var tikt iegūtas tuvinātas vērtības, kuru precizitāte ir tikai viens cipars aiz komata, divi cipari, un līdz pat 16 cipariem aiz komata. Tādēļ ir nepieciešams apskatīt, kā mainās zaudējumu funkcijas: saknes vidējās kvadrātiskās kļūdas (*RMSE*) vērtības starp novēroto un prognozēto *NDVI* laika rindu atkarībā no aproksimētā apakšsignāla precizitātes.

Katrai no 100 *NDVI* laika rindām katrai novērojumu izlasei (70 % no visiem novērojumiem ik pa vienu laika soli uz priekšu) trešā eksperimenta laikā tiek atrasts tāds prognozēšanas modelis, un datu pirmapstrādes parametri, kas dod minimālo *RMSE* vērtību uz validācijas datu kopas, izmantojot apakšsignālu ar 16 cipariem aiz komata. Tālāk tiek veikts ceturtais eksperiments, kura **mērķis** ir pārbaudīt, kā mainās *RMSE* vērtības atkarībā no apakšsignāla precizitātes. Viss eksperimenta apraksts sakrīt ar trešā eksperimenta aprakstu, izņemot to, ka apakšignāliem secīgi tiek atmests noteikts skaits pēdējo ciparu aiz komata. Respektīvi, eksperiments tiek veikts ar apakšsignāliem, kuriem ir 16 cipari aiz komata, tad 15 cipari aiz komata, un visbeidzot, kuriem ir tikai viens cipars aiz komata. Un pie katra ciparu skaita aiz komata tiek aprēķināta minimāla *RMSE* vērtība uz validācijas datu kopas.

Tad atrastais piemērotais prognozēšanās modelis un datu pirmapstrādes parametri tiek izmantoti, lai prognozētu vienu vērtību uz priekšu katrai novērojumu izlasei. Rezultātā visas prognozētas vērtības no katras novērojumu izlases izveido testa datu kopu, uz kuras tiek aprēķināta *RMSE* vērtība starp novēroto un prognozēto *NDVI* laika rindu testa kopai atbilstošajos laika soļos, un tad visām 100 laika rindām tiek aprēķināta visu *RMSE* vērtību vidējā vērtība.

3.3.3. Rezultātu analīze eksperimentiem ar modificēto VMD metodi

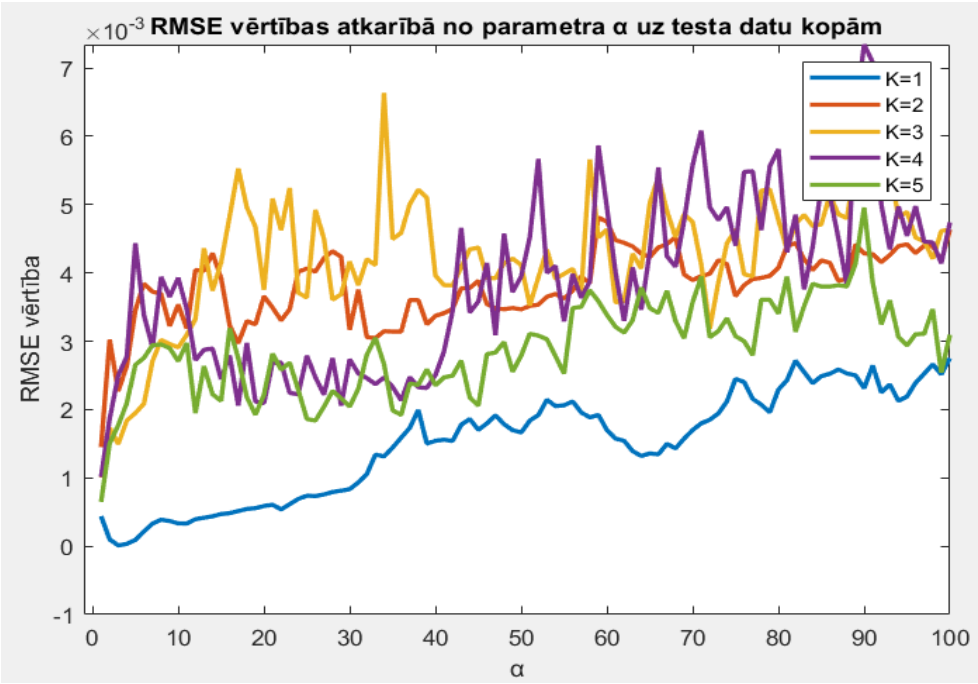
Trešajā eksperimentā iegūtas saknes vidējās kvadrātiskās kļūdas *RMSE*, virzienu simetrijas *DS* un determinācijas koeficienta R^2 vidējās vērtības no visām 100 *NDVI* laika rindām uz visām datu kopām ir redzamas 3.4. tabulā.

3.4. tabula

Eksperimentā ar modificēto *VMD* metodi iegūtās zaudējumu funkciju vērtības

Datu kopa	*RMSE*	*DS*	R^2
Apmācības datu kopa	0,000058	99,88%	1,00
Validācijas datu kopa	0,000076	100%	1,00
Testa datu kopa	0,000074	100%	0,99

Saknes vidējās kvadrātiskās kļūdas vērtības atkarībā no apakšsignālu summas precizitāti regulējošā parametra α pie dažāda apakšsignālu skaita K redzamas 3.4. attēlā.



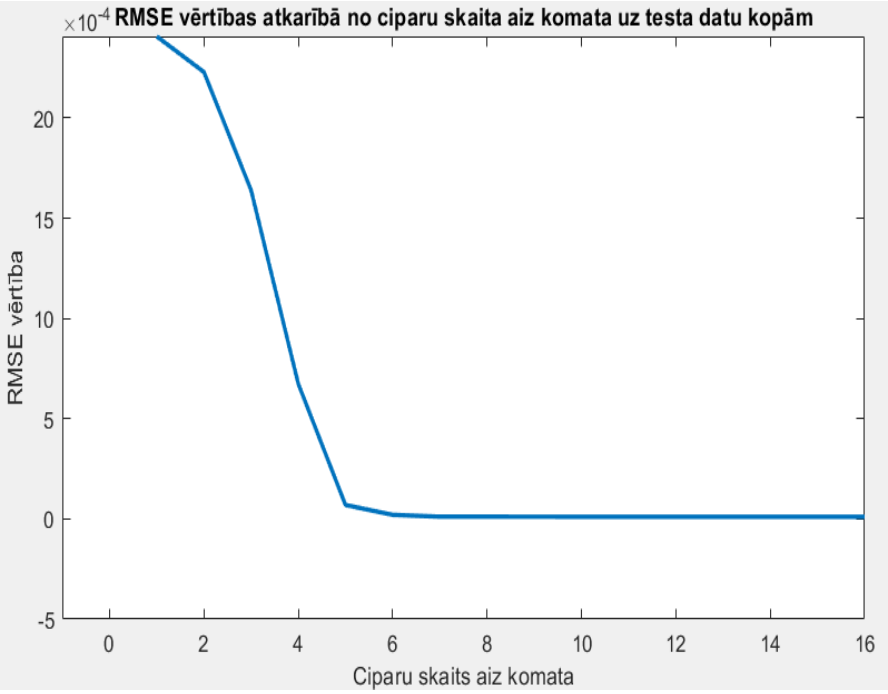
3.4. att. *RMSE* vērtības atkarībā no precizitātes parametra α.

Kā redzams 3.4. attēlā, pirmkārt, *RMSE* vērtībām ir tendence pieaugt, pieaugot apakšsignālu summas precizitāti regulējošā parametra α vērtībām, lai gan šis pieaugums nav monotons, un ir sastopami gan *RMSE* vērtību kāpumi, gan kritumi pie

dažādām α vērtībām. Otrkārt, ir redzams, ka viszemākās *RMSE* vērtības tiek sasniegtas, izmantojot tikai vienu apakšsignālu. Tas nozīmē, ka, izmantojot autora piedāvāto modificēto *VMD* metodi, ir pilnīgi pietiekami veikt *NDVI* laika rindu dekompozīciju ar vienu apakšsignālu un apakšsignālu summas precizitāti regulējošā parametra α vērtības meklēt intervālā [0,01; 10] ar soli 0,01.

Jaunu laika rindas vērtību (vienas vērtības uz priekšu katrā novērojumu izlasē) prognozēšana pēc trešā eksperimenta uzrāda augstu precizitāti. Tomēr oriģinālās laika rindas reducēšanas dēļ, prognozēšana ir iespējama tikai oriģinālās *NDVI* laika rindas vidusdaļā. Lai varētu prognozēt arī jaunas laika rindas vērtības, kas atrodas aiz visiem vēsturiskajiem novērojumiem, ir nepieciešams izstrādāt pieeju, kas ļauj tuvināti aprēķināt apakšsignāla vērtības, neizmantojot pašu modificēto *VMD* metodi katrā laika solī. Šāda aproksimācijas pieejas izstrādāšana tiek veikta 3.4. apakšnodaļā.

Ceturtajā eksperimentā iegūtas 16 vidējās *RMSE* vērtības pie dažāda apakšsignāla ciparu skaita aiz komata, kas ir dotas 3.5. attēlā.



3.5. att. *RMSE* vērtības atkarībā no apakšsignāla ciparu skaita aiz komata.

Kā var redzēt 3.5. attēlā, tad *RMSE* vērtības strauji samazinās no viena līdz pieciem cipariem aiz komata. No sešiem cipariem līdz visiem 16 cipariem aiz komata *RMSE* vērtību izmaiņas ir nelielas. No viena līdz trim cipariem aiz komata vidējā *RMSE* vērtība svārstās starp 0,00164 un 0,0024. Ar četriem cipariem aiz komata tiek iegūta vidējā *RMSE* vērtība 0,00067. Savukārt, sākot no pieciem

ciparіem, un līdz pat 16 ciparіem aiz komata, vidējā *RMSE* vērtība svārstās starp 0,00049 un 0,000074.

Pirmajā eksperimentā tiek izmantota oriģinālā *VMD* metode, to pielietojot visiem *NDVI* laika rindas vēsturiskajiem novērojumiem. No 100 *NDVI* laika rindām uz validācijas datu kopām tiek iegūta vidējā *RMSE* vērtība 0,000040, bet uz testa datu kopām tiek iegūta vidējā *RMSE* vērtība 0,0460. Tātad uz validācijas datu kopām precizitāte, kas tiek iegūta, prognozējot *NDVI* laika rindas, izmantojot no oriģinālās un modificētās *VMD* metodes iegūtos apakšsignālus, ir līdzvērtīga. Šeit apakšsignāli ir iegūti, padodot gan oriģinālajai, gan modificētajai *VMD* metodei ieejā visus *NDVI* laika rindas vēsturiskos novērojumus. Bet uz testa datu kopām modificētās *VMD* metodes izmantošana *NDVI* laika rindu prognozēšanas uzdevumā spēj uzlabot sasniegto prognozēšanas precizitāti.

Otrajā eksperimentā tiek izmantota oriģinālā *VMD* metode, to lietojot katram *NDVI* laika rindas fragmentam, lai veiktu tā dekompozīciju, un iegūtie apakšsignāli tiek izmantoti kā ieejas dati prognozēšanas modelī. No 100 *NDVI* laika rindām uz testa datu kopām tiek iegūta vidējā *RMSE* vērtība 0,0011. Tādēļ, lai modificētā *VMD* metode, to izmantojot kopā ar apakšsignāla aproksimācijas pieeju, sniegtu augstāku *NDVI* laika rindu prognozēšanas precizitāti, nekā, izmantojot oriģinālo *VMD* metodi katram *NDVI* laika rindas fragmentam, aproksimētajām apakšsignālam tuvināti būtu jāiegūst vismaz četri pareizi cipari aiz komata, salīdzinot ar apakšsignālu, ko izejā dod modificētā *VMD* metode.

3.4. Apakšsignāla vērtību tuvinātā aprēķināšana

Izmantojot laika rindu prognozēšanai modificēto *VMD* metodi, iegūtais apakšsignāla novērojumu skaits ir aptuveni divas reizes mazāks nekā oriģinālās *NDVI* laika rindas novērojumu skaits. Turklāt apakšsignāla novērojumi atbilst oriģinālās *NDVI* laika rindas vēsturisko novērojumu vidusdaļai. Lai varētu veikt *NDVI* laika rindu prognozēšanu, par atribūtiem izmantojot apakšsignāla vērtības, respektīvi, meklēt funkcionālo sakarību (1.7.), ir nepieciešams izstrādāt pieeju, kas ļauj tuvināti aprēķināt apakšsignāla vērtības jebkurā laika solī t. Šim nolūkam autors piedāvā izmantot lineāru vienādojumu sistēmu (*System of Linear Equations, LES*), risināšanu. Lineāro vienādojumu sistēmu [55] apraksta vienādojums (3.4.):

$$Aw = b, \tag{3.4.}$$

kur A – koeficientu matrica (mxn);

b – brīvo locekļu vektors ($mx1$);

w – mainīgo vektors ($nx1$);

m – vienādojumu skaits;

n – nezināmo skaits.

Atrisinot *LES* (skat. vienādojumu (3.4.)) katrā laika solī t, tiktu iegūti n nezināmie, kur no šiem nezināmajiem varētu iegūt tuvinātu apakšsignāla vērtību laika solī t. Lai varētu iegūt *LES* unikālu risinājumu, ir jāizpildās šādiem nosacījumiem [56]:

1) matricai A ir jābūt kvadrātiskai ($m = n$);
2) matricas A rangam ir jābūt vienādam ar n (visām matricas A kolonnām, tāpat kā rindām, ir jābūt lineāri neatkarīgām).

Ja šie nosacījumi izpildās, tad *LES* (3.4.) unikālu atrisinājumu w [22] iegūst pēc izteiksmes (3.5.):

$$w = A^{-1}b. \tag{3.5.}$$

kur A^{-1} – koeficientu matricas A inversā matrica.

Jāņem vērā arī koeficientu matricas A stabilitāte, ko nosaka nosacītības skaitlis (*condition number*). Matricas A nosacītības skaitli [55] aprēķina pēc izteiksmes (3.6.):

$$cond\left(A\right) = \frac{\lambda_{max}}{\lambda_{min}}, \tag{3.6.}$$

kur $\lambda_{\min}$ – matricas A mazākā īpašvērtība;

$\lambda_{\max}$ – matricas A lielākā īpašvērtība.

Ja nosacītības skaitlis ir liels, tad matrica A ir slikti nosacīta un nelielas kļūdas sistēmas (3.4.) vektorā b var izsaukt lielas kļūdas atrisinājumu vektorā w. Nosacītības skaitlis var pieņemt vērtības no 1 līdz bezgalībai. Jo lielāks paliek šis skaitlis, jo pieaug kļūdas lielums *LES* risinājumā. Tātad koeficientu matricas A nosacītības skaitlim ir jābūt pietiekami mazam, lai varētu iegūt atrisinājumu, kur vismaz sakrīt četri vai vairāk cipari aiz komata, salīdzinot ar oriģinālo apakšsignālu.

Izstrādātajā pieejā vispirms tiek atrisināta homogēnā lineāro vienādojumu sistēma, iegūstot fundamentālu atrisinājumu sistēmu, ko veido lineāri neatkarīgo atrisinājumu kopa. Tad tiek izstrādāta pirmā nehomogēnā lineāro vienādojumu sistēma, kuras partikulārais atrisinājums tiek kombinēts ar lineāri neatkarīgo homogēnās lineāro vienādojumu sistēmas atrisinājumu kopu, iegūstot koeficientu matricu, kas tiek izmantota otrajā nehomogēnajā lineāro vienādojumu sistēmā. Otrās nehomogēnās lineāro vienādojumu sistēmas partikulārais atrisinājums tiek atrasts katrā laika solī t, atbilstoši mainot brīvo locekļu vektoru b.

3.4.1. Homogēnās lineāro vienādojumu sistēmas atrisinājums

Ja lineāro vienādojumu sistēmas (*LES*) visi brīvie koeficienti ir vienādi ar nulli (b ir nuļļu vektors), tad sistēmu sauc par homogēnu [10], ko aprēķina pēc (3.7.):

$$Ax = 0 . \tag{3.7.}$$

Homogēnās *LES* (skat. vienādojumu sistēmu (3.7.)) atrisinājumu kopu x sauc par matricas A fundamentālu atrisinājumu sistēmu [10]. Risinot homogēno *LES,* ir iespējami divi varianti, cik tiks iegūti lineāri neatkarīgi atrisinājumi:

1) ja koeficientu matricas A lineāri neatkarīgo rindu skaits vienādojumā (3.7.) ir vienāds ar tās nezināmo (kolonnu) skaitu, tad var iegūt tikai triviālu atrisinājumu. Tātad lineāri neatkarīgu atrisinājumu skaits ir nulle;
2) ja koeficientu matricas A lineāri neatkarīgu rindu skaits ir mazāks nekā tās nezināmo skaits, tad fundamentālu atrisinājumu sistēmas lineāri neatkarīgo atrisinājumu skaits ir vienāds ar starpību starp matricas A lineāri neatkarīgo rindu skaitu un nezināmo skaitu.

Ja oriģinālās *NDVI* laika rindas novērojumu skaits $T = 570$, tad, padodot šo laika rindu ieejā modificētajai *VMD* metodei ar apakšsignālu skaitu $K = 1$, tiek iegūts viens apakšsignāls u ar 285 novērojumiem ($N = 285$). Tātad, izvēloties atbilstošus atribūtus un izveidojot kvadrātisko koeficientu matricu A, ir iespējams iegūt homogēnās *LES* (skat. vienādojumu sistēmu (3.7.)) atrisinājumu kopu ar maksimums $N-1$ lineāri neatkarīgiem atrisinājumiem un N mainīgajiem katrā risinājumā. Ir nepieciešams izveidot tādu homogēnu lineāro vienādojumu sistēmu (skat. vienādojumu sistēmu (3.7)), kur vismaz viens nezināmais no atrisinājumu vektora x būtu no modificētās *VMD* metodes iegūtā apakšsignāla u vērtība noteiktā laika solī

(nezināmais var būt arī transformētā veidā) un, atrisinot homogēnu *LES*, tiktu iegūta fundamentālu atrisinājumu sistēma ar $N-1$ lineāri neatkarīgiem atrisinājumiem un N mainīgajiem katrā atrisinājumā.

Lai iegūtu kvadrātisku koeficientu matricu A, var tikt izmantotas mašīnmācīšanā izmantotās kodola funkcijas (*kernel functions*). Kodola funkcija K nodrošina datu transformāciju no ievades telpas uz daudzu dimensiju atribūtu jeb pazīmju telpu [74]. Vienkāršākā kodola funkcija ir lineārā kodola funkcija [25], ko apraksta izteiksme (3.8.):

$$K(X,Y) = XY^T + c, \tag{3.8.}$$

kur X – ieejas datu kopa;

Y^T – transponēta ieejas datu kopa Y;

c – konstante.

Ja ieejas datu kopu veido apakšsignāls u, kura novērojumu skaits ir N, tad, padodot šo apakšsignālu ieejā lineārajai kodola funkcijai ar konstanti $c=0$, tiek iegūta NxN matrica, un to apraksta ar matricu (3.9.):

$$K(u,u) = \begin{bmatrix} u(1)u(1) & u(1)u(2) & \dots & u(1)u(N) \\ u(2)u(1) & u(2)u(2) & \dots & u(2)u(N) \\ \dots & \dots & & \dots \\ u(N)u(1) & u(N)u(2) & \dots & u(N)u(N) \end{bmatrix}. \tag{3.9.}$$

Gadījumā, ja $N=285$ (kā darbā izmantotajiem datiem), iegūtā matrica (skat. matricu (3.9.)) būs $285x285$ liela matrica. Par homogēnā lineāro vienādojumu sistēmā (skat. vienādojumu sistēmu (3.7.)) izmantoto koeficientu matricu A var tikt izmantota matrica $K(u,u)$ (skat. matricu (3.9.)), un, atrisinot šo *LES* ar singulāro vērtību dekompozīcijas (*Singular Value Decomposition, SVD*) metodi, tiek iegūta fundamentālu atrisinājumu sistēma x, kur šo lineāri neatkarīgo atrisinājumu kopu x apraksta matrica (3.10.):

$$x = \begin{bmatrix} x_1(1) & x_1(2) & x_1(3) & \dots & x_1(N) \\ x_2(1) & x_2(2) & x_2(3) & \dots & x_2(N) \\ \dots & \dots & \dots & & \dots \\ x_{N-1}(1) & x_{N-1}(2) & x_{N-1}(3) & \dots & x_{N-1}(N) \end{bmatrix}, \tag{3.10.}$$

kuru veido $N-1$ lineāri neatkarīgie atrisinājumi un N mainīgie. Katrs atrisinājums no

fundamentālu atrisinājumu sistēmas (skat. matricu (3.10.)) – tas ir svaru vektors, kuru, reizinot ar jebkuru matricas (3.9.) rindu (skalārais vektoru reizinājums), tiek iegūta nulle, kā parādīts vienādojumā (3.11.):

$$K_i(1)x_j(1) \quad + \quad K_i(2)x_j(2) \quad + \quad ... \quad K_i(N)x_j(N) = 0, \qquad (3.11.)$$

kur $\boldsymbol{i}$ – lineārās kodola funkcijas (3.8.) rindu indekss, $\boldsymbol{i = 1, \ldots, N}$;

$\boldsymbol{j}$ – fundamentālu atrisinājumu sistēmas (3.10.) atrisinājumu indekss, $j = 1, ..., N-1$.

Tagad var pārrakstīt lineārā kodola funkciju (3.9.) kā lineārā kodola funkciju, kas ieejā saņem datu kopu, kas sastāv no diviem vektoriem y un u, kā tas ir parādīts matricā (3.12.):

$$K([yu],[yu]) = \begin{bmatrix} y(1)y(1)+u(1)u(1) & \ldots & y(1)y(N)+u(1)u(N) \\ y(2)y(1)+u(2)u(1) & \ldots & y(2)y(N)+u(2)u(N) \\ \ldots & & \ldots \\ y(N)y(1)+u(N)u(1) & \ldots & y(N)y(N)+u(N)u(N) \end{bmatrix}. \qquad (3.12.)$$

Matricas (3.12.) jebkuru rindu reizinot ar jebkuru atrisinājumu (skalārais vektoru reizinājums) no fundamentālu atrisinājumu sistēmas (3.10.), saskaitot visus locekļus, kas satur apakšsignāla u vērtības, iegūst nulli. Bet, saskaitot visus locekļus, kas satur oriģinālās laika rindas y vērtības, tiek iegūta kāda summa, kas sastāda brīvo locekļu vektora atbilstošās rindas vērtību.

3.4.2. Nehomogēno lineāro vienādojumu sistēmu atrisinājumi

Lai varētu katrā laika solī t tuvināti aprēķināt apakšsignāla vērtību $u(t)$, izmantojot lineāru vienādojumu sistēmu (3.4.), ir nepieciešama kvadrātiska koeficientu matrica A. Tomēr fundamentālu atrisinājumu sistēmas (3.10.) atrisinājumu kopā ir par vienu atrisinājumu mazāk nekā nezināmo skaits. Ir nepieciešams iegūt vēl vienu atrisinājumu – atrast pirmās nehomogēnās lineāro vienādojumu sistēmas partikulāru risinājumu w, kur kā koeficientu matrica A_1 var tikt izmantota matrica (3.13.) ar izmēriem $NxN-1$. Matricas (3.13.) pirmo atribūtu veido vienīnieku vektors, bet, sākot no otrā atribūta līdz pirmspēdējam atribūtam, ir lineārās kodola funkcijas (3.12.) atbilstošie atribūti. Savukārt pēdējais atribūts ir *NDVI* laika rinda y.

$$A_1 = \begin{bmatrix} 1 & y(1)y(1)+u(1)u(1) & \dots & y(1)y(N-2)+u(1)u(N-2) & y(1) \\ 1 & y(2)y(1)+u(2)u(1) & \dots & y(2)y(N-2)+u(2)u(N-2) & y(2) \\ \dots & & \dots & \dots & \dots \\ 1 & y(N)y(1)+u(N)u(1) & \dots & y(N)y(N-2)+u(N)u(N-2) & y(N) \end{bmatrix}. \qquad (3.13.)$$

Savukārt par brīvo locekļu vektoru b tiek piedāvāts izmantot *NDVI* laika rindu y. Lai iegūtu pirmās nehomogēnās lineāro vienādojumu sistēmas partikulāro atrisinājumu w, tiek izmantota nelineāra regresijas analīze ar mazāko kvadrātu metodi. Kā nelineāra funkcija f tiek izmantota funkcija (3.14.):

$$f = w(1) + \frac{\sqrt{S^2+1}-1}{2} + S + yw(N), \qquad (3.14.)$$

kur S – svērtā summa.

Svērtās summas elementu $S(t)$, kur $t = 1,...,N$ aprēķinā pēc (3.15.) vienādojuma:

$$S(t) = w(2) + \sum_{i=2}^{N-2} A_1(t,i)w(j),\ j = 3,...,N-1. \qquad (3.15.)$$

Pievienojot partikulāro atrisinājumu w pie fundamentālu atrisinājumu sistēmas, tiek iegūta koeficientu matrica A_2, kā parādīts izteiksmē (3.16.):

$$A_2 = \begin{bmatrix} x_1(1) & x_1(2) & \dots & x_1(N-2) & x_1(N-1) & x_1(N) \\ x_2(1) & x_2(2) & \dots & x_2(N-2) & x_2(N-1) & x_2(N) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{N-1}(1) & x_{N-1}(2) & \dots & x_{N-1}(N-2) & x_{N-1}(N-1) & x_{N-1}(N) \\ w(3) & w(4) & \dots & 0 & 0 & 0 \end{bmatrix}. \qquad (3.16.)$$

Iegūtā koeficientu matrica A_2 ir kvadrātiska ar N risinājumiem un N nezināmajiem, un tā tiek izmantota otrajā nehomogēnajā lineāro vienādojumu sistēmā. Brīvo locekļu vektors b, kas tiek izmantots otrajā nehomogēnajā *LES*, attiecīgi arī sastāv no N elementiem. Vektora $b_t(i)$ elementi laika solī t, kur $i = 1,\dots,N-1$ tiek iegūti, izmantojot izteiksmi (3.17.):

$$b_t(i)=\begin{bmatrix} y(t)y(1)x_1(1) & + & y(t)y(2)x_1(2) & + & \dots & + & y(t)y(N)x_1(N) \\ y(t)y(1)x_2(1) & + & y(t)y(2)x_2(2) & + & \dots & + & y(t)y(N)x_2(N) \\ \dots & \dots & \dots & & \dots & \dots & \dots & \dots \\ y(t)y(1)x_{N-1}(1) & + & y(t)y(2)x_{N-1}(2) & + & \dots & + & y(t)y(N)x_{N-1}(N) \end{bmatrix}. \qquad (3.17.)$$

Izmantojot partikulāro atrisinājumu w un brīvo locekļu vektoru b, var tikt aprēķināta nezināma svērta summa $\hat{S}$. Un vektora $b_t(i)$ elements, kur $i=N$, tad tiek aprēķināts pēc vienādojuma (3.18.):

$$b_t(i)=\hat{S}(t)-w(2). \qquad (3.18.)$$

Tādējādi katrā laika solī t, atrisinot otro nehomogēno lineāro vienādojumu sistēmu pēc (3.5.), kur koeficientu matrica A_2 tiek iegūta pēc izteiksmes (3.16.), un brīvo locekļu vektors b_t pēc izteiksmes (3.17.) un vienādojuma (3.18.), tiek iegūts otrās nehomogēnās lineāro vienādojumu sistēmas partikulārais atrisinājums w_t (skat. izteiksmi (3.19.)), kurš sastāv no N mainīgajiem:

$$w_t=\begin{bmatrix} w_t(1) & w_t(2) & \dots & w_t(N) \end{bmatrix}. \qquad (3.19.)$$

Lai iegūtu apakšsignāla u aproksimēto vērtību laika solī t, vispirms tiek izmantota izteiksme (3.20.), kas ļauj iegūt N dažādas $\hat{u}_i(t)$ vērtības, kur $i=1,\dots,N$:

$$\hat{u}_i(t)=\frac{\left(w_t(i)-y(t)y(i)\right)}{u(i)}. \qquad (3.20.)$$

Un visiem i tiek aprēķināta vidējā vērtība pēc izteiksmes (3.21.):

$$\hat{u}(t)=\frac{\sum_{i=1}^{N}\hat{u}(t)_i}{N}. \qquad (3.21.)$$

Pēc izteiksmes (3.21.) iegūtais rezultāts ir apakšsignāla u aproksimētā vērtība laika solī t. Atkarībā no otrajā nehomogēnajā lineāro vienādojumu sistēmā izmantotās koeficientu matricas A_2 (skat. matricu (3.16.)) nosacītības skaitļa, iegūtās aproksimētās vērtības var būt ar dažādu precizitāti – ar dažādu ciparu skaitu, kas sakrīt ar oriģinālā apakšsignāla u cipariem. Izmantojot koeficientu matricas A_2 nosacītības skaitli $cond(A_2)$ var tikt aprēķināts aptuvenais ciparu skaits M (izteiksme (3.22.)),

pa kādu samazinās precizitāte, iegūstot lineāro vienādojumu sistēmas atrisinājumu [100]:

$$M = \log_{10}(cond(A_2)). \tag{3.22.}$$

Otrs faktors, kas ietekmē aproksimētā apakšsignāla precizitāti, ir lineāro vienādojumu sistēmu atrisinājumu kļūdu lielumi. Salīdzināt oriģinālo un aproksimēto apakšsignālu gan var tikai tajos laika soļos, kuros ir pieejams oriģinālais apakšsignāls, respektīvi, *NDVI* laika rindas vēsturisko novērojumu vidusdaļai atbilstošajos laika soļos.

Šīs aproksimācijas pieejas izmantošanai tiek lietota četrkārtīga precizitāte (*quadruple precision*), kas aprēķinos izmanto 34 ciparus aiz komata. Aproksimētā apakšsignāla $\hat{u}$ pienesumu *NDVI* laika rindu prognozēšanas uzdevumā kopumā var novērtēt tikai izmantojot to kā ieejas datus prognozēšanas modelim un tad, novērtējot iegūto prognozēšanas precizitāti, kas tiek veikts 5. nodaļā.

3.5. Kopsavilkums un secinājumi

Nodaļa veltīta variāciju apakšsignālu dekompozīcijas (*VMD*) metodes modifikācijas izstrādāšanai un no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieejas izstrādāšanai, kas ļauj tuvināti katrā laika solī aprēķināt apakšsignāla vērtības. Tiek veikti četri eksperimenti, kur tiek veikta *NDVI* laika rindu prognozēšana, izmantojot no oriģinālās un modificētās variāciju apakšsignālu dekompozīcijas metodes iegūtos apakšsignālus. Galvenie secinājumi par veiktajiem eksperimentiem, izstrādāto *VMD* metodes modifikāciju un no šīs metodes iegūtā apakšsignāla aproksimācijas pieeju ir šādi:

1. Prognozējot *NDVI* laika rindas, oriģinālo variāciju apakšsignāla dekompozīcijas metodi, ja tās ieejā tiek padoti visi vēsturiskie laika rindas novērojumi, ir jāizmanto katrā laika solī, jo laika rindai nāk klāt jauni novērojumi, un šajos laika soļos arī ir jābūt aprēķinātām atbilstošām apakšsignālu vērtībām. Tomēr katru reizi izmantojot *VMD* metodi, kad *NDVI* laika rindai nāk klāt jauns novērojums, visas iepriekš aprēķinātās apakšsignālu vērtības mainās, un apmācītie prognozēšanas modeļi, kuri izmantoja šos apakšsignālus kā ieejas datus, vairs nav lietojami.
2. Ja, prognozējot *NDVI* laika rindas, oriģinālajai variāciju apakšsignāla dekompozīcijas metodei ieejā tiek padoti *NDVI* laika rindas fragmenti ar kādu

noteiktu loga garumu, tad arī oriģinālo *VMD* metodi ir jāizmanto katrā laika solī, bet iepriekš aprēķinātās apakšsignālu vērtības nemainās. Tādēļ apmācītus prognozēšanas modeļus, kuri izmanto šos apakšsignālus kā ieejas datus, var izmantot arī jaunu *NDVI* vērtību prognozēšanai.

3. Prognozējot *NDVI* laika rindu, kur oriģinālā *VMD* metode tiek pielietota visiem vēsturiskajiem *NDVI* laika rindas novērojumiem, tiek konstatēts, ka augstu prognozēšanas precizitāti var sasniegt uz apmācības un validācijas datu kopām, bet uz testa datu kopām precizitāte strauji krīt.
4. Prognozējot *NDVI* laika rindu, kur oriģinālā *VMD* metode tiek pielietota katram *NDVI* laika rindas fragmentam atsevišķi, tiek konstatēts, ka uz visām trim datu kopām tiek sasniegta līdzvērtīga precizitāte.
5. Analizējot oriģinālās *VMD* metodes pielietošanu *NDVI* laika rindas prognozēšanas uzdevumā, tiek konstatēts, ka variāciju apakšsignālu dekompozīcijas metodes autori izmanto "spoguļošanu" [118], kas divreiz pagarina ieejas laika rindu, un iegūtie apakšsignāli satur informāciju par paplašinātās laika rindas vērtībām.
6. Ja oriģinālajai *VMD* metodei ieejā tiek padoti visi vēsturiskie *NDVI* laika rindas novērojumi, tad iegūtos apakšsignālus padodot ieejā prognozēšanas modelim, tas apmācās tā, ka, pielietojot to jaunas *NDVI* laika rindas vērtības prognozēšanai laika solī t, kā prognoze tiek izdota vērtība, kas ir tuva paplašinātās laika rindas vērtībai šajā laika solī.
7. Risinot "spoguļošanas" problēmu, autors piedāvā *VMD* metodes modifikāciju, kur netiek izmantota "spoguļošana", bet divreiz tiek samazināts novērojumu skaits, ko ieejā saņem *VMD* metode, turklāt ieejā tiek saņemta vēsturisko laika rindas novērojumu vidusdaļa, kamēr abās laika rindas malās esošie novērojumi kalpo laika rindas paplašināšanai, aizvietojot "spoguļošanu". Tas ļauj izmantot modificēto *VMD* metodi arī vienas nākamās laika rindas vērtības prognozēšanai, kas pēc laika atrodas aiz modificētās *VMD* metodes ieejas datiem, respektīvi, vienu laika soli aiz *NDVI* laika rindas vidusdaļas novērojumiem.
8. Modificētā *VMD* metode izejā dod apakšsignālus, kuri atbilst vēsturisko *NDVI* laika rindas novērojumu vidusdaļas laika soļiem, tādēļ ir nepieciešama pieeja, kas ļauj tuvināti aprēķināt apakšsignāla vērtības jebkurā laika solī, kur ir pieejami vēsturiskie *NDVI* laika rindas novērojumi.

9. Kā risinājumu problēmai ar modificētās *VMD* metodes izmantošanu jaunu *NDVI* laika rindas vērtību prognozēšanai, autors piedāvā pieeju, kas, izmantojot no modificētās *VMD* metodes iegūtā apakšsignāla vērtības, vispirms, atrisinot vairākas lineāro vienādojumu sistēmas, iegūst koeficientu matricu, un, tad katrā laika solī mainot brīvo locekļu vektoru, tiek atrisināta lineāro vienādojumu sistēma, kuras atrisinājums ir tuvināta apakšsignāla vērtība šajā laika solī.

Izstrādātas apakšsignāla aproksimācijas pieejas izmantošanai tiek lietota četrkārtīga precizitāte, kas aprēķinos izmanto precizitāti ar 34 cipariem aiz komata.

4. PROGNOZĒŠANAS SISTĒMAS *NDVI FS* IZSTRĀDĀŠANA

Normalizētais veģetācijas indekss (*Normalized Difference Vegetation Index, NDVI*) plaši tiek izmantots ar veģetācijas dinamiku saistītajās nozarēs, piemēram, lauksaimniecībā, mežsaimniecībā un apkārtējās vides aizsardzībā. *NDVI* indeksa laika rindas ir viens no galvenajiem instrumentiem precīzajā lauksamniecībā, kas palīdz sekot līdzi dažādu lauksaimniecības kultūru augšanas tempiem un regulēt barības vielu izlietojuma apjomu. Šī indeksa prognozēšana precīzajā lauksaimniecībā ļauj ātri noteikt problēmas, kas saistītas ar lauksaimniecības kultūru augšanu, un savlaicīgi pieņemt lēmumus par nepieciešamajiem pasākumiem, lai novērstu šīs problēmas.

Latvijā publiskā vidē pieejamie pētījumi, kas veltīti *NDVI* indeksa laika rindu prognozēšanai, tāpat kā gatavi produkti lietojumprogrammas veidā *NDVI* indeksa prognozēšanai, nav izplatīti. Tāpēc ir vitāli svarīgi izstrādāt *NDVI* laika rindu īstermiņa prognozēšanas sistēmu, kuras izmantošana spētu nodrošināt precīzās lauksaimniecības nozares attīstību Latvijā. *NDVI* indeksa prognozēšanas sistēma (*NDVI Forecasting System, NDVI FS*) risina *NDVI* laika rindu analīzi un prognozēšanu ar datu pirmapstrādes, signālu apstrādes un prognozēšanas metodēm, atbilstoši sistēmas konceptuālai shēmai, kā parādīts 1.5. attēlā.

4.1. Sistēmas *NDVI FS* funkcijas

Normalizētā veģetācijas indeksa (*NDVI*) laika rindu prognozēšanas sistēmas (*NDVI FS*) ietvaros tiek piedāvāts risinājums precīzās lauksaimniecības nozares speciālistiem, kā efektīvāk apstrādāt *NDVI* laika rindas ar datu pirmapstrādes metodēm un pieejām, lai tālāk varētu veikt precīzāku īstermiņa *NDVI* indeksa prognozēšanu. Lai izveidotu *NDVI FS*, ir jārealizē vairākas funkcijas.

1. Vēsturisko *NDVI* datu apkopošana. Sistēmai lietotāja izvēlētajam pikselim ir jāizveido skalāra *NDVI* laika rinda, izmantojot datu krātuvē esošos vēsturiskos *NDVI* attēlus.
2. Sagatavoto datu pirmapstrāde. Sistēma veic *VMD* apakšsignāla aproksimāciju (ja lietotājs ir izvēlējies izmantot aproksimācijas pieeju), iegūst *NDVI* laika rindas un aproksimētā apakšsignāla rekonstruētās fāžu telpas (ja lietotājs nav izvēlējies *VMD* apakšsignāla aproksimācijas pieeju, tad tikai *NDVI* laika rindas fāžu telpa tiek rekonstruēta), veic fāžu telpu apvienošanu (ja ir divas fāžu

telpas). Tiek veikta apmācības, validācijas un testa datu kopu izveidošana. Informatīvāko atribūtu atlasīšana ļauj reducēt izmantojamo atribūtu skaitu. Apmācības, validācijas un testa atribūtu kopas tiek transformētas uz lineāri nekorelējošu atribūtu kopām.

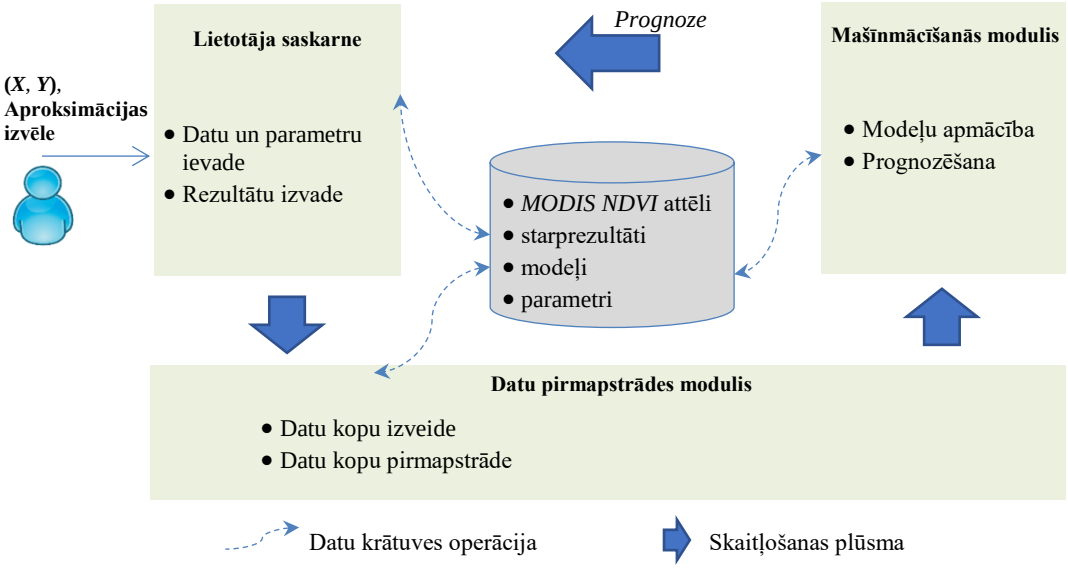
3. Prognozēšanas modeļa apmācība. Apmācības un validācijas datu kopas pēc pirmapstrādes veikšanas tiek padotas ieejā rekurento slāņu neironu tīklam, notiek apmācība, un tiek iegūts prognozēšanas modelis.
4. Izveidotās sistēmas precizitātes pārbaude. Iegūtais prognozēšanas modelis tiek pielietots, lai īstermiņā prognozētu vēsturiskās *NDVI* laika rindas vērtības vienu soli uz priekšu. Zaudējumu funkciju vērtības tiek aprēķinātas uz apmācības, validācijas un testa datu kopām.
5. Jaunu *NDVI* vērtību prognozēšana. Iegūtais prognozēšanas modelis tiek izmantots jaunu *NDVI* laika rindas vērtību īstermiņa prognozēšanai vienu soli uz priekšu.
6. Iegūto rezultātu atspoguļošana. Lietotājam tiek parādīta iegūtā prognoze, kā arī aprēķinātās zaudējumu funkciju vērtības vēsturiskajos datos (apmācības, validācijas un testa datu kopās).

NDVI laika rindu prognozēšanas sistēma pamatojas uz iepriekšējo nedēļu *NDVI* vērtībām – vēsturiskajiem datiem.

4.2. Sistēmas *NDVI FS* vispārīgā arhitektūra un darbības princips

Prognozēšanas sistēma *NDVI FS* sastāv no lietotāja saskarnes, datu pirmapstrādes moduļa, mašīnmācīšanās moduļa un datu krātuves. Prognozēšanas sistēmas *NDVI FS* vispārīgā arhitektūra ir parādīta 4.1. attēlā.

Izmantojot lietotāja saskarni, sistēmā tiek ievadīti visi nepieciešamie dati un parametri, kas nodrošina tās darbību. Tāpat izmantojot šo saskarni, lietotājam tiek izvadīta *NDVI* īstermiņa prognoze, kas arī ir sistēmas izeja. Datu pirmapstrādes modulī tiek izveidotas datu kopas un tiek veikta to pirmapstrāde. Pirmapstrādātas datu kopas nonāk mašīnmācīšanās modulī, kur tiek veikta prognozēšanas modeļu apmācība un īstermiņa prognozēšana. Savukārt datu krātuvē tiek glabāti vidējās izšķirtspējas attēlveidošanas spektroradiometra (*Moderate Resolution Imaging Spectroradiometer, MODIS*) *NDVI* attēli, dažādi starprezultāti, prognozēšanas modeļi un datu pirmapstrādes parametri.



4.1. att. Prognozēšanas sistēmas *NDVI FS* vispārīgā arhitektūra.

Uzsākot darbu, lietotājs, izmantojot atbilstošo saskarni, ievada sistēmā trīs nepieciešamo elementu vērtības:

- Koordināta X jeb ģeogrāfiskais garums *MODIS NDVI* attēlos atbilstoši šo attēlu koordinātu sistēmai.
- Koordināta Y jeb ģeogrāfiskais platums *MODIS NDVI* attēlos atbilstoši šo attēlu koordinātu sistēmai.
- izvēles elementa vērtība (1: izmantot *VMD* apakšsignāla aproksimācijas pieeju, 0: neizmantot).

Pēc koordinātu un izvēles elementa vērtību saņemšanas, sistēma izveido vaicājumu, kurš tiek nosūtīts datu krātuvei. Ja konkrētā pikseļa (atbilstoši ievadītājām koordinātām) laika rindai un atbilstošajai izvēles elementa vērtībai eksistē prognozēšanas modelis kopā ar piemērotiem datu pirmapstrādes parametriem, tad tie tiek atlasīti vaicājumos.

Ja konkrētā pikseļa laika rindai un atbilstošajai izvēles elementa vērtībai neeksistē prognozēšanas modelis kopā ar piemērotiem datu pirmapstrādes parametriem, tad analizējamajā datu kopā, kas sastāv no šīs *NDVI* laika rindas, tiek veikts datu pirmapstrādes process (skat. 4.3. apakšnodaļu).

- Vaicājumā no datu krātuves tiek atlasīti visi pieejamie *MODIS NDVI* attēli.
- No atlasītajiem attēliem konkrētajam pikselim, atbilstoši lietotāja ievadītajām koordinātām, tiek iegūts datu masīvs, kas satur *NDVI* vēsturiskās vērtības.

- Datu masīvs tālāk tiek saglabāts ar komatiem atdalīto vērtību (*Comma-Separated Values, CSV*) datnē.
- No *CSV* datnes sistēmā tiek ielasīti visi dati, iegūstot skalāru *NDVI* laika rindu, un šī laika rinda tiek padota datu pirmapstrādes modulim.

NDVI laika rindai tiek veikta dekompozīcija, ja lietotājs ir izvēlējies šādu datu pirmapstrādes posmu, izmantojot darba autora piedāvātu no modificētās variāciju apakšsignālu dekompozīcijas (*Variational Mode Decomposition, VMD*) metodes iegūtā apakšsignāla aproksimācijas pieeju.

- Dekompozīcijas rezultātā iegūtajai laika rindai tiek veikta fāžu telpas rekonstrukcija ar laika aiztures metodi.
- Arī oriģinālajai *NDVI* laika rindai tiek veikta fāžu telpas rekonstrukcija ar laika aiztures metodi,
- Abas fāžu telpas tiek apvienotas vienotā datu kopā, un šī kopa tiek sadalīta apmācības (70 % no apvienotas datu kopas ierakstiem), validācijas (15 % no datu ierakstiem) un testa (15 % no datu ierakstiem) datu kopās.
- Katra no šīm trim kopām tiek sadalīta atribūtu kopā, kuru veido katras datu kopas (apmācības, validācijas vai testa) ieraksti no pirmā līdz pirmspēdējam ierakstam, un prognozējamajā parametrā, kuru veido katras datu kopas ieraksti no otrā līdz pēdējam ierakstam.

Ja lietotājs nav izvēlējies izmantot *VMD* apakšsignāla aproksimācijas pieeju, tad tiek rekonstruēta tikai viena fāžu telpa, kuru iegūst no oriģinālās *NDVI* laika rindas, un šī fāžu telpa tiek sadalīta apmācības (70 % no datu ierakstiem), validācijas (15 % no datu ierakstiem) un testa (15 % no datu ierakstiem) datu kopās.

- Katra no šīm trim kopām tiek sadalīta atribūtu kopā, kuru veido katras datu kopas ieraksti no pirmā līdz pirmspēdējam ierakstam, un prognozējamajā parametrā, kuru veido katras datu kopas ieraksti no otrā līdz pēdējam ierakstam.
- Visām trim atribūtu kopām tālāk tiek veikta informatīvāko atribūtu atlasīšana, izmantojot pakāpenisko regresijas analīzi, un atribūtu kopu transformācija uz lineāri nekorelējošu atribūtu kopām, izmantojot principiālo komponentu analīzi.
- Rezultātā tiek iegūtas reducētas, lineāri nekorelējošu atribūtu kopas (apmācības, validācijas un testa atribūtu kopas).

Piemērotās datu pirmapstrādes metožu parametru vērtības tiek saglabātas datu krātuvē. Datu kopas, kas tiek iegūtas pēc datu pirmapstrādes procesa pabeigšanas, saņem ieejā rekurento slāņu neironu tīkls (*Layer Recurrent Neural Network*, *LRNN*). Tiek veikta apmācība, *LRNN* svaru un konstanšu vērtības, kas ļauj iegūt konkrētās *NDVI* laika rindas precīzākas prognozes, izveido prognozēšanas modeli, kas savukārt arī tiek saglabāts datu krātuvē.

Savukārt, ja konkrētā pikseļa laika rindai un atbilstošajai izvēles elementa vērtībai eksistē prognozēšanas modelis kopā ar piemērotiem datu pirmapstrādes parametriem, tad ar vaicājumu no datu krātuves tiek atlasīti pēdējie 50 attēli. Konkrētā pikseļa laika rindas vērtības tiek apstrādātas ar atbilstošām datu pirmapstrādes metodēm, kuru parametru vērtības tiek atlasītas no datu krātuves ar vaicājumiem. Tālāk tiek veikta prognozēšana, izmantojot ar vaicājumu atlasītu *LRNN* prognozēšanas modeli.

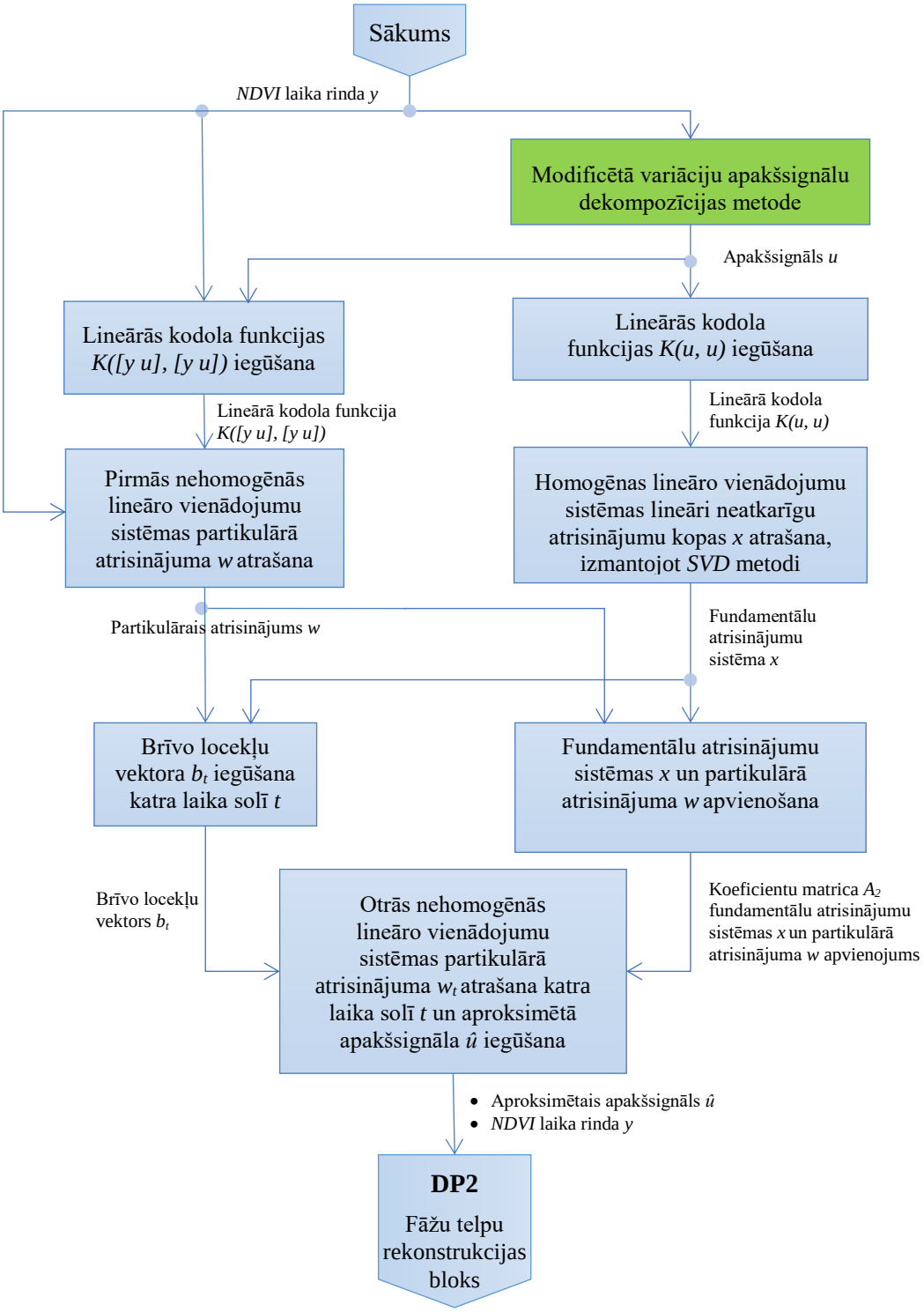
4.3. Datu pirmapstrādes modulis

Ja lietotājs ir ievadījis izvēles elementa vērtību viens, un tiek izmantota *VMD* apakšsignāla aproksimācijas pieeja, tad, izmantojot bloku "DP1", kurš ietver *VMD* apakšsignāla aproksimācijas algoritmu, tiek iegūta jauna laika rinda – apakšsignāls. Ja lietotājs ir ievadījis izvēles elementa vērtību nulle, tad blokā "DP1" netiek veikta apakšsignāla aproksimācija, un datu pirmapstrāde sākas no bloka "DP2".

4.3.1. Apakšsignāla aproksimācija

Bloks "DP1" ieejā saņem *NDVI* laika rindu y. Šī laika rinda tiek padota ieejā modificētajai *VMD* metodei, kura ir aprakstīta 3.2. apakšnodaļā. Modificētā *VMD* metode izejā dod apakšsignālu. Pēc apakšsignāla iegūšanas tiek veikta tā aproksimācija, izmantojot algoritmu, kas aprakstīts 3.4. apakšnodaļā, un parādīts 4.2. attēlā.

- Modificētā *VMD* metode izejā dod apakšsignālu u, kas tiek izmantots lineārās kodola funkcijas $K(u,u)$ izveidošanai, kura ir aprakstītas matricā (3.9.).
- Iegūto lineāro kodola funkciju izmanto kā koeficientu matricu A homogēnā lineāro vienādojumu sistēmā.



4.2. att. Apakšsignāla aproksimācijas bloks “DP1”.

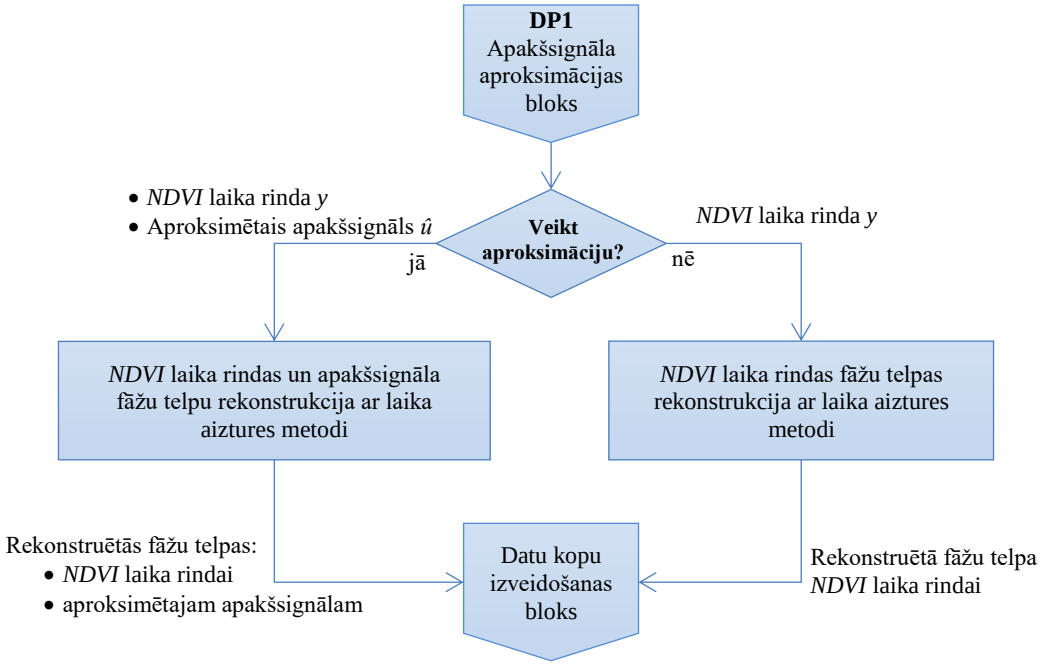
Atrisinot homogēno lineāro vienādojumu sistēmu ar singulāro vērtību dekompozīcijas (*SVD*) metodi, iegūst lineāri neatkarīgu atrisinājumu kopu – fundamentālu atrisinājumu sistēmu x, ko apraksta matrica (3.10.).

- Tālāk tiek iegūta otrā lineārā kodola funkcija $K([yu],[yu])$, izmantojot apakšsignālu u un *NDVI* laika rindu y (skat. matricu (3.12.)).
- Pirmās nehomogēnās lineāro vienādojumu sistēmas partikulārais atrisinājums w tiek atrasts, izmantojot mazāko kvadrātu metodi, kur izteiksme (3.13.) tiek izmantota kā koeficientu matrica A_1, bet *NDVI* laika rinda y tiek izmantota kā brīvo locekļu vektors b_1.
- No šī partikulārā atrisinājuma tiek paņemti atbilstošie regresijas koeficienti, un tie tiek apvienoti ar fundamentālu atrisinājumu sistēmu (3.10.), iegūstot koeficientu matricu A_2 (3.16.), kuru izmanto otrajā nehomogēnajā lineāro vienādojumu sistēmā. Savukārt brīvo locekļu vektoru b_t aprēķina pēc vienādojumu sistēmas (3.17.) un vienādojuma (3.18.).
- Atrodot otrās nehomogēnās lineāro vienādojumu sistēmas partikulāru atrisinājumu w_t (skat. izteiksmi (3.19.)) laika solī t, iegūst aproksimētā apakšsignāla $\hat{u}$ vērtību laika solī t pēc izteiksmēm (3.19.) un (3.21.). Aproksimētais apakšsignāls $\hat{u}(t)$, kur $t = 1,\ldots,N$, un N ir *NDVI* laika rindas y vēsturisko novērojumu skaits, arī ir bloka "DP1" izeja.

Blokā "DP1" iegūta otrās nehomogēnās lineāro vienādojumu sistēmas koeficientu matrica un pirmās nehomogēnās lineāro vienādojumu sistēmas partikulārais atrisinājums ar vaicājuma palīdzību tiek ierakstīti datu krātuvē, kur unikālais identifikators sasaista to ar konkrēto *NDVI* laika rindu. Bloka "DP1" izeja tiek padota ieejā nākamajam blokam "DP2".

4.3.2. Fāžu telpu rekonstrukcija

Fāžu telpu rekonstrukcijas blokā (bloks "DP2") tiek veikta katras laika rindas fāžu telpas PhS rekonstrukcija ar laika aiztures metodi atbilstoši vienādojumam (2.23). Kā fāžu telpas dimensija m tiek izmantota vērtība 50, savukārt laika aiztures τ vērtība ir viens. Fāžu telpas dimensija arī tiek izmantota kā loga garums, pēc kura gan *NDVI* laika rinda, gan aproksimētais apakšsignāls (ja lietotājs ir izvēlējies veikt dekompozīciju un tuvināti aprēķināt apakšsignālu) tiek sadalīti fragmentos. Fāžu telpu rekonstrukcijas bloka "DP2" skaitļošanas plūsma parādīta 4.3. attēlā.



4.3. att. Fāžu telpu rekonstrukcijas bloks "DP2".

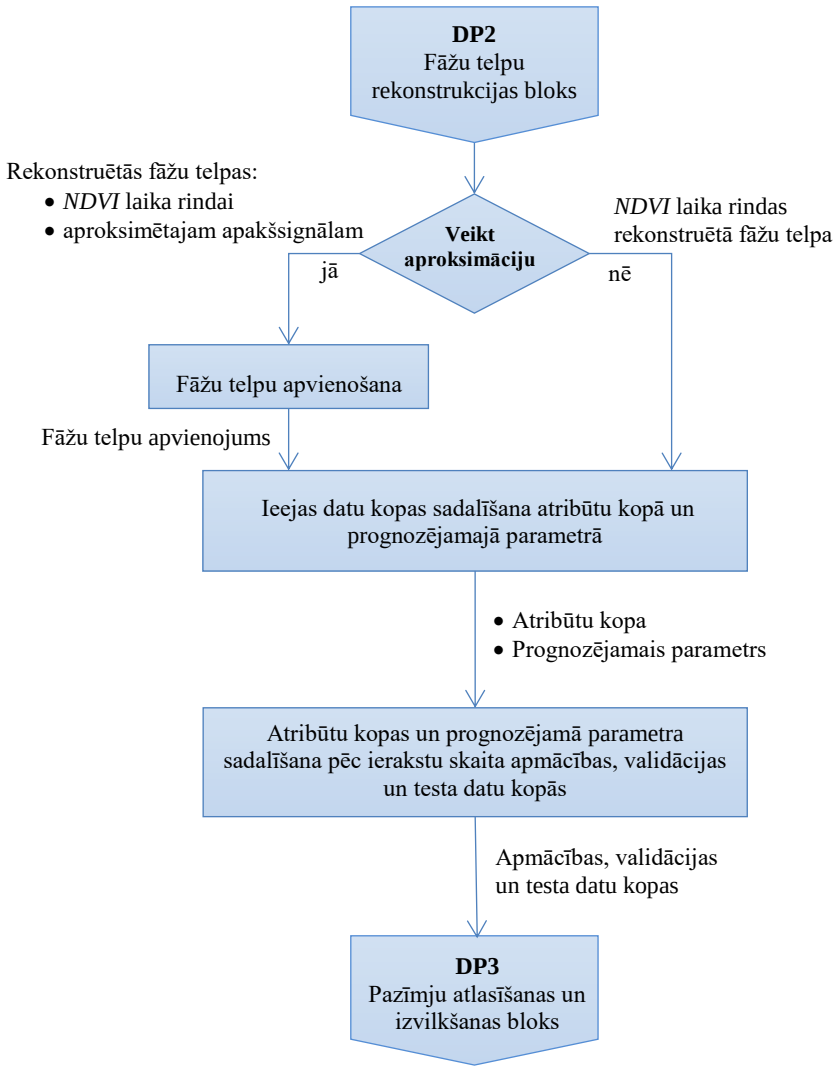
Ja lietotājs ir ievadījis no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieejas izvēles elementa vērtību nulle, tad tikai *NDVI* laika rindai tiek veikta fāžu telpas *PhS* rekonstrukcija ar laika aiztures metodi, un rekonstruētā fāžu telpa tālāk tiek padota ieejā datu kopu izveidošanas blokam.

Savukārt, ja tiek izmantota *VMD* apakšsignāla aproksimācijas pieeja, tad divas fāžu telpas tālāk tiek padotas datu kopu izveidošanas blokam.

4.3.3. Datu kopu izveidošana

Datu kopu izveidošanas blokā tiek apvienotas fāžu telpas (ja ieejā tiek saņemtas divas fāžu telpas), vai arī tiek izmantota tikai viena fāžu telpa (ja netiek izmantota apakšsignāla aproksimācijas pieeja) un tiek iegūta datu kopa. Skaitļošanas plūsma datu kopu izveidošanas blokā parādīta 4.4. attēlā.

- Iegūtā datu kopa tiek sadalīta divās daļās: atribūtu jeb neatkarīgo parametru kopā, kuru veido iegūtās datu kopas ieraksti no pirmā līdz pirmspēdējam ierakstam, un atkarīgajā jeb prognozējamajā parametrā, kuru veido no *NDVI* laika rindas rekonstruētās fāžu telpas ieraksti no otrā līdz pēdējām ierakstam.



4.4. att. Datu kopu izveidošanas bloks.

- Gan atribūtu kopa, gan prognozējamais parametrs tiek sadalīti trīs daļās, iegūstot apmācības (70 % no datu ierakstiem), validācijas (15 % no datu ierakstiem) un testa (15 % no datu ierakstiem) datu kopas. Tā tiek sagatavotas datu kopas pakāpeniskās regresijas analīzei, principiālo komponentu analīzei un rekurento slāņu neironu tīklu apmācībai.
- Ja *VMD* apakšsignāla aproksimācijas pieeja netiek izmantota, tad tikai *NDVI* laika rindas fāžu telpa tiek padota ieejā datu kopu izveidošanas blokam.

Apmācības, validācijas un testa datu kopas tiek padotas pēdējam datu pirmapstrādes blokam “DP3”.

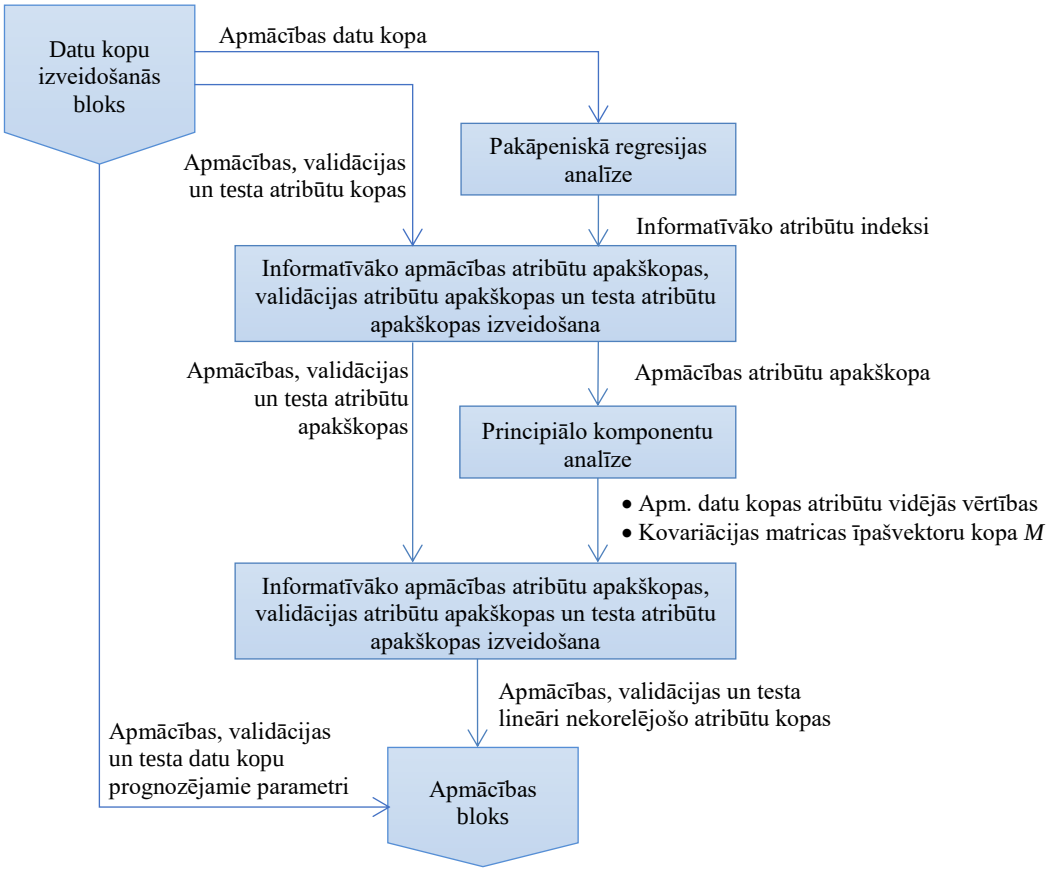
4.3.4. Pazīmju atlasīšana un izvilkšana

Pazīmju atlasīšanas un izvilkšanas blokā (bloks "DP3") tiek veikta informatīvāko atribūtu atlasīšana (neinformatīvākie atribūti tiek izslēgti no atribūtu kopām), izmantojot pakāpenisko regresijas analīzi (*Stepwise Regression Analysis*), un lineāri neatkarīgu atribūtu iegūšana, izmantojot principiālo komponentu analīzi (*Principal Component Analysis, PCA*).

- Izmantojot pakāpenisko regresijas analīzi, apmācības datu kopas atkarīgais parametrs tiek prognozēts, ņemot dažādas atribūtu apakškopas no apmācības atribūtu kopas, atbilstoši pakāpeniskās regresijas analīzes algoritmam (skat. 2.3. apakšnodaļu).
- Kad no apmācības atribūtu kopas ir iegūta informatīvāko atribūtu apakškopa, kas, prognozējot atkarīgo parametru, ļauj iegūt minimālu *RMSE* vērtību, tad šādas atribūtu apakškopas ar iepriekš atrastajiem informatīvāko atribūtu indeksiem tiek izveidotas arī no validācijas un testa atribūtu kopām.
- Tālāk šīs atlasītās atribūtu apakškopas tiek transformētas par lineāri nekorelējošu atribūtu kopām, veicot principiālo komponentu analīzi (skat. 2.4. apakšnodaļu).
- Principiālo komponentu analīzes metode ieejā saņem apmācības atribūtu kopu, un tiek iegūtas nepieciešamās parametru vērtības: apmācības atribūtu vidējās vērtības un apmācības atribūtu kovariācijas matricas C īpašvektoru kopa M, kas ļauj iegūt no apmācības atribūtu kopas lineāri nekorelējošu atribūtu kopu. Tad šie parametri tiek izmantoti, lai arī no validācijas un testa atribūtu kopām iegūtu lineāri nekorelējošu atribūtu kopas.

Skaitļošanas plūsma blokā "DP3" – informatīvāko atribūtu atlasīšana, izmantojot pakāpenisko regresijas analīzi un reducēto atribūtu apakškopu transformācija, izmantojot principiālo komponentu analīzi, parādīta 4.5. attēlā. Tagad iegūtās lineāri nekorelējošu apmācības, validācijas un testa atribūtu kopas kopā ar apmācības, validācijas un testa datu kopu prognozējamo parametru ir datu pirmapstrādes moduļa izeja. Visi nepieciešamie datu pirmapstrādes parametri, kas ir nepieciešami atribūtu atlasīšanai, izmantojot pakāpenisko regresijas analīzi, un atribūtu transformācijai par lineāri nekorelējošiem atribūtiem, izmantojot principiālo komponentu analīzi, ir šādi:

- informatīvāko atribūtu indeksi (kolonnu indeksi atribūtu kopās);
- apmācības datu kopas atribūtu vidējās vērtības;
- apmācības datu kopas atribūtu kovariācijas matricas C īpašvektoru kopa M .



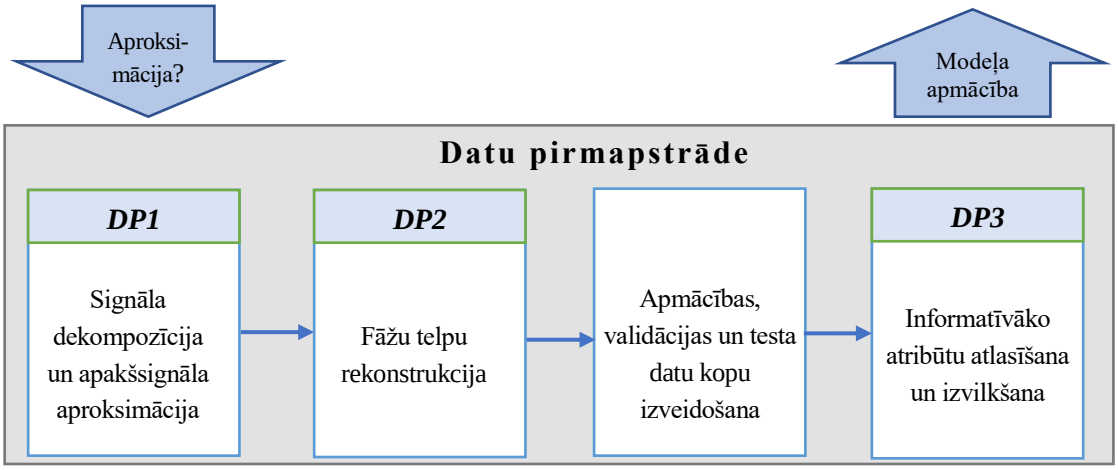
4.5. att. Pazīmju atlasīšanas un izvilkšanas bloks “DP3”.

Tie tiek iegūti, apstrādājot konkrēto *NDVI* laika rindu, tiek saglabāti datu krātuvē, ar unikālo identifikatoru, sasaistot tos gan ar konkrēto *NDVI* laika rindu, gan ar iepriekš saglabāto otrās nehomogēnās lineāro vienādojumu sistēmas koeficientu matricu un pirmās nehomogēnās lineāro vienādojumu sistēmas partikulāro atrisinājumu.

4.3.5. Pirmapstrādes moduļa skaitļošanas plūsma

Ņemot vērā augstāk aprakstītus pirmapstrādes soļus un bloku struktūru, datu pirmapstrādes moduļa skaitļošanas plūsma (4.6. attēls) ir šāda. Modulis ieejā saņem *NDVI* laika rindu, kas iegūta no pikseļa, ko ir izvēlējies lietotājs, ievadot koordinātas. Tāpat modulis ieejā saņem aproksimācijas pieejas izvēles parametra vērtību. Ja lietotājs ir izvēlējies izmantot apakšsignāla aproksimācijas pieeju, tad datu pirmapstrādes

process sākas ar aproksimācijas bloku “DP1”, kur no modificētas *VMD* metodes tiek iegūts apakšsignāls, un tiek veikta tā aproksimācija. Aproksimētais apakšsignāls un *NDVI* laika rinda tiek padoti ieejā fāžu telpu rekonstrukcijas blokam “DP1”, kur tiek veikta abu laika rindu fāžu telpu rekonstrukcija. Ja lietotājs nav izvēlējies izmantot apakšsignāla aproksimācijas pieeju, tad datu pirmapstrādes modulī skaitļošanas plūsma sākas ar fāžu telpu rekonstrukcijas bloku, kurš ieejā saņem *NDVI* laika rindu un veic tās fāžu telpas rekonstrukciju. Divas (vai viena) fāžu telpas tālāk tiek padotas datu kopu izveidošanas blokam, kur tiek iegūtas apmācības, validācijas un testa datu kopas, un tās tiek padotas ieejā pazīmju atlasīšanas un izvilkšanas blokam “DP3”. Pazīmju atlasīšanas un izvilkšanas bloka izejā iegūtās pirmapstrādātās apmācības, validācijas un testa datu kopas tiek padotas ieejā apmācības blokam.



4.6. att. Datu pirmapstrādes modulis.

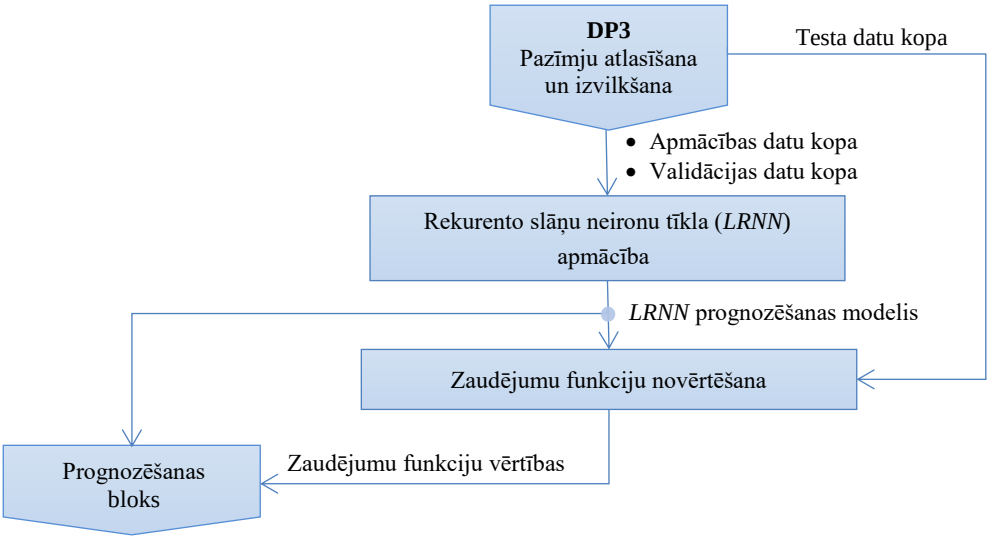
Datu pirmapstrādes moduļa izejā ir pirmapstrādātās apmācības, validācijas un testa datu kopas, kas nonāk ieejā mašīnmācīšanās modulī. Datu pirmapstrādes gaitā iegūtie parametri: koeficientu matrica A_2; partikulārais atrisinājums w; informatīvāko atribūtu indeksi; atribūtu vidējās vērtības un īpašvektoru kopa M tiek saglabāti datu krātuvē, ar unikālo identifikatoru sasaistot tos ar konkrēto *NDVI* laika rindu.

4.4. Mašīnmācīšanās modulis

Mašīnmācīšanās moduli veido divas komponentes: apmācības bloks un prognozēšanas bloks. Ieejā modulis saņem pirmapstrādātās apmācības, validācijas un testa datu kopas no datu pirmapstrādes moduļa. Savukārt moduļa izejā ir iegūta prognoze, kas tiek izsniegta lietotājam.

4.4.1. *Prognozēšanas modeļa apmācība*

Apmācības blokā tiek iegūts rekurento slāņu neironu tīkla (*LRNN*) prognozēšanas modelis. Šīs prognozēšanas modelis tiek iegūts uz apmācības datu kopas, novērtējot *RMSE* vērtību starp novēroto un prognozēto *NDVI* laika rindu uz validācijas datu kopas. Kā piemērotākais *LRNN* prognozēšanas modelis tiek izvēlēts tas, kuram tiek sasniegta mazākā zaudējumu funkcijas *RMSE* vērtība. Tad apmācītais prognozēšanas modelis tiek izmantots, lai novērtētu zaudējumu funkciju: *RMSE*, *DS* un pielāgota determinācijas koeficienta R^2_{adj} vērtības uz apmācības, validācijas un testa datu kopām. Skaitļošanas plūsma apmācības blokā parādīta 4.7. attēlā.



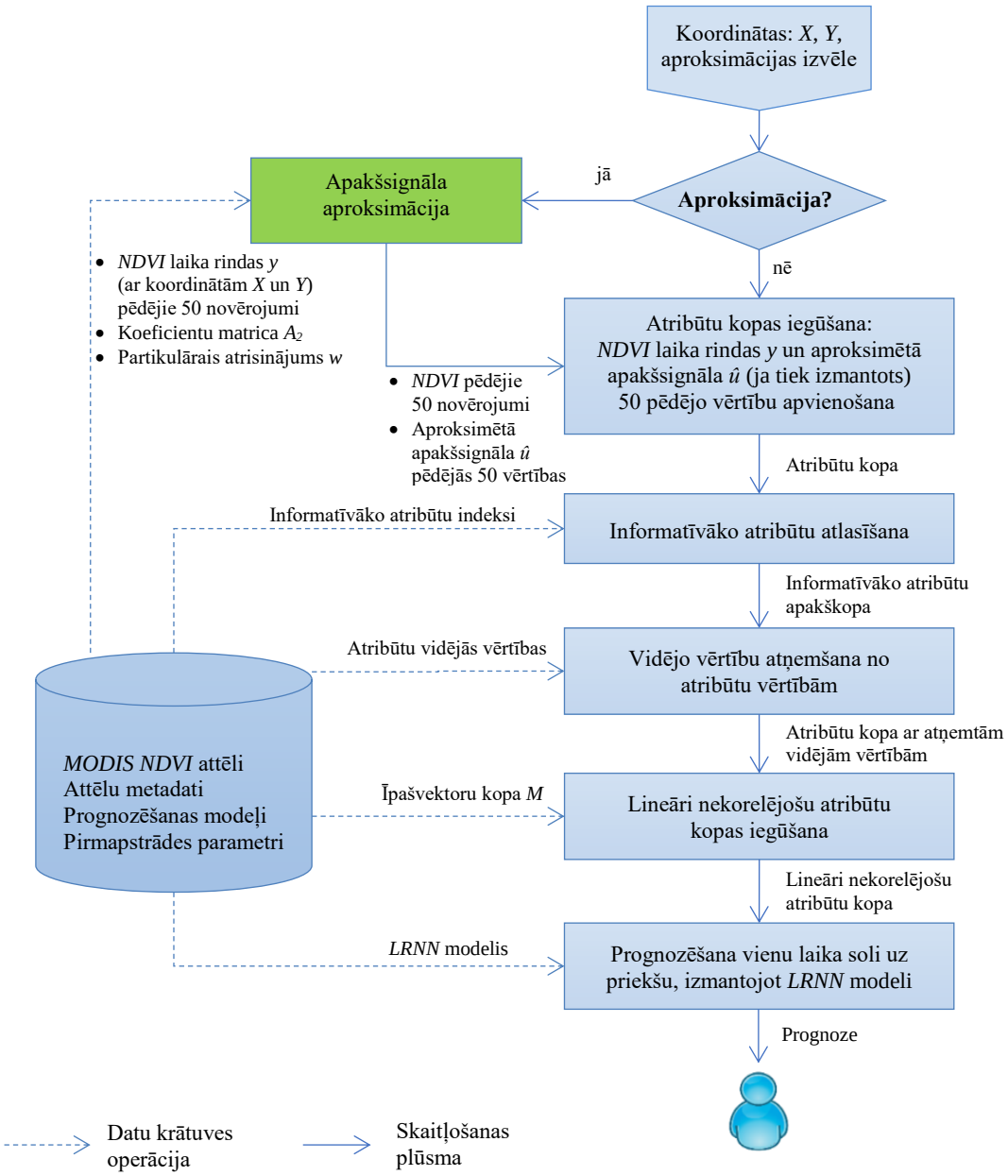
4.7. att. Mākslīgo neironu tīklu apmācības bloks.

Kad tiek atrasts piemērotākais *LRNN* prognozēšanas modelis, tad šī *LRNN* modeļa parametri (svari un konstantes) tiek saglabāti datu krātuvē, tādējādi iegūstot konkrētajai *NDVI* laika rindai piemērotāko prognozēšanas modeli. Apmācības bloka izejā ir prognozēšanas modelis.

4.4.2. *NDVI laika rindas prognozēšana*

Kad prognozēšanas modelis ir iegūts, to var izmantot *NDVI* laika rindas vērtību prognozēšanai, to veicot *NDVI* laika rindu prognozēšanas blokā (skat. 4.8. attēlu).

- No datu krātuves ar vaicājumiem tiek iegūti pēdējie 50 *MODIS NDVI* attēli, no kuriem konkrētajam pikselim atbilstošajai skalārajai *NDVI* laika rindai y iegūst pēdējos 50 novērojumus.



4.8. att. Prognozēšanas bloks.

- Šie novērojumi iziet datu pirmapstrādes procesu ar datu krātuvē jau saglabātiem parametriem.
- Ja konkrētajai *NDVI* laika rindai y, veicot datu pirmapstrādi, tika arī izmantota *VMD* apakšsignāla aproksimācijas pieeja, tad vispirms, izmantojot saglabāto koeficientu matricu A_2 un partikulāro atrisinājumu w, 50 reizes ar dažādiem brīvo locekļu vektoriem b_t tiek atrisināta otrā nehomogēnā lineāro vienādojumu sistēma, iegūstot 50 pēdējās aproksimētā apakšsignāla $\hat{u}$ vērtības.
- No *NDVI* laika rindas y 50 pēdējiem novērojumiem un aproksimētā

apakšsignāla $\hat{u}$ (ja tas tiek izmantots) 50 pēdējiem novērojumiem tiek atlasīti tie novērojumi, kuri atbilst informatīvāko atribūtu indeksiem.

- No atlasītajiem informatīvākajiem atribūtiem tiek atņemtas vidējās vērtības un datu ieraksts, kas sastāv no šo atribūtu vērtībām (rindas vektors) tālāk tiek reizinātas ar īpašvektoru kopu M (matricu), iegūstot transformētu atribūtu vērtību kopu.
- Pirmapstrādāto atribūtu kopa tiek padota *LRNN* prognozēšanas modelim ieejā. Izmantojot *LRNN* prognozēšanas modeļa parametrus, tiek veikta prognozēšana vienam laika solim uz priekšu. Iegūtā prognoze tiek izvadīta lietotājam.

Pēc noteikta laika perioda, kas *MODIS NDVI* attēlu gadījumā ir septiņas dienas, 4.8. attēlā parādītais prognozēšanas process var tikt atkārtots, prognozējot jau nākamo vērtību.

4.5. Izstrādātā prognozēšanas sistēma *NDVI FS*

Darbā izstrādāta prognozēšanas sistēma *NDVI FS*. Tā paredzēta normalizētā veģetācijas indeksa (*NDVI*) laika rindu īstermiņa prognozēšanai. Sistēma *NDVI FS* ieejā saņem *NDVI* laika rindas vēsturiskos novērojumus, bet izejā izsniedz prognozi. Prognozēšanas sistēmas *NDVI FS* arhitektūra, datu krātuves operācijas un skaitļošanas plūsma ir parādīta 4.9. attēlā.

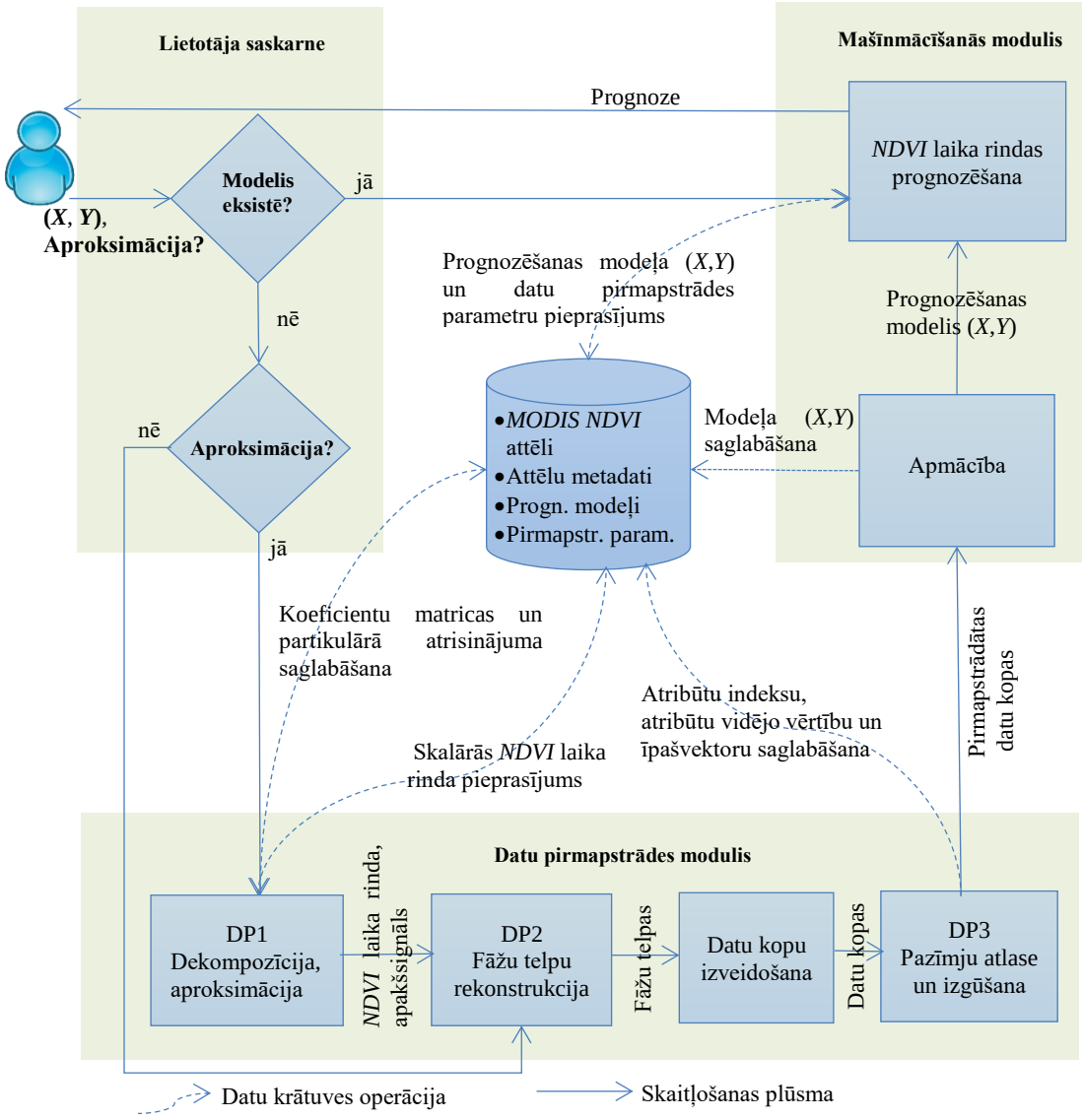
4.5.1. Sistēmas NDVI FS arhitektūra un skaitļošanas plūsma

Prognozēšanas sistēma *NDVI FS* sastāv no trim komponentēm – lietotāja saskarnes, datu pirmapstrādes moduļa un mašīnmācīšanās moduļa. Informācija par *MODIS NDVI* attēliem – attēla nosaukums, metadati, paši attēli, iepriekš atrastas datu pirmapstrādes parametru vērtības un rekurento slāņu neironu tīkla (*LRNN*) prognozēšanas modeļi, kā arī dažādi starprezultāti glabājas datu krātuvē.

Lietotāja saskarne atbild par pikseļa koordināšu ievadi un no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieejas izvēli (izmantot vai neizmantot). Tāpat caur lietotāja saskarni tiek izvadīta prognoze, kas tiek iegūta sistēmas izejā.

Izmantojot vaicājumu datu krātuvei, tiek pārbaudīts, vai jau eksistē prognozēšanas modelis un datu pirmapstrādes parametri lietotāja izvēlētai *NDVI* laika rindai. Ja tie eksistē, tad skaitļošanas plūsma turpinās mašīnmācīšanās moduļa

prognozēšanas blokā. No datu krātuves ar vaicājumiem tiek atlasīti pēdējie 50 *MODIS NDVI* attēli, atbilstošais prognozēšanas modelis un datu pirmapstrādes parametri.



4.9. att. *NDVI* laika rindu prognozēšanas sistēmas *NDVI FS* arhitektūra.

No pēdējiem 50 *MODIS NDVI* attēliem tiek atlasīta tā pikseļa laika rinda, ko ir ievadījis lietotājs. Šī laika rinda iziet datu pirmapstrādes procesu ar iepriekš atlasītām parametru vērtībām, kas ir atlasītas no datu krātuves, un pirmapstrādātas datu kopas tiek padotas ieejā atlasītajām *LRNN* modelim, kas sniedz prognozi, kura savukārt tiek attēlota lietotājam, izmantojot atbilstošo saskarni.

Ja prognozēšanas modelis un datu pirmapstrādes parametri neeksistē lietotāja izvēlētā pikseļa laika rindai, vai, ja tie eksistē šai laika rindai, bet nesakrīt aproksimācijas pieejas izmantošanas izvēli, tad skaitļošanas plūsma nonāk datu

pirmapstrādes modulī, aproksimācijas blokā “DP1”. Vispirms no datu krātuves ar vaicājumu tiek atlasīti visi pieejami *MODIS NDVI* attēli, un tiek iegūta skalāra *NDVI* laika rinda lietotāja ievadītajām koordinātām. Ja lietotājs ir izvēlējies izmantot apakšsignāla aproksimāciju, tad blokā “DP1” tiek veikta laika rindas dekompozīcija un iegūtā apakšsignāla aproksimācija, un abas laika rindas tiek padotas ieejā fāžu telpu rekonstrukcijas blokam “DP2”. Ja lietotājs nav izvēlējies izmantot apakšsignāla aproksimācijas pieeju, tad bloks “DP2” ieejā saņem tikai *NDVI* laika rindu. Blokā “DP2” tiek rekonstruētas fāžu telpas abām laika rindām (vai tikai vienai, ja lietotājs nav izvēlējies izmantot aproksimācijas pieeju). Abas fāžu telpas (vai arī tikai viena fāžu telpa) tiek padotas ieejā datu kopu izveidošanas blokam, kur tiek iegūtas apmācības, validācijas un testa datu kopas. Šīs datu kopas tiek padotas ieejā pazīmju atlasīšanas un izvilkšanas blokam “DP3”. Blokā “DP3” no visām trim datu kopām tiek atlasīti informatīvākie atribūti, un veikta to transformācija, lai iegūtu lineāri nekorelējošus atribūtus. Pirmapstrādātas datu kopas nonāk mašīnmācīšanās modulī, apmācības bloka ieejā. Apmācības blokā tiek veikta prognozēšanas modeļa apmācība. Iegūtais prognozēšanas modelis nonāk ieejā prognozēšanas blokam, kurā tiek veikta *NDVI* laika rindas īstermiņa prognozēšana. Prognoze tiek izsniegta lietotājam caur atbilstošo saskarni. Visi nepieciešamie datu pirmapstrādes parametri un prognozēšanas modelis tiek saglabāti datu krātuvē, izveidojot mapi ar nosaukumu, kas satur pikseļa numuru pēc kārtas *MODIS NDVI* attēlos. Numurs tiek iegūts, secīgi skaitot pikseļus no pirmās rindas un pirmās kolonnas pikseļa pa labi līdz pēdējam pikselim pirmajā rindā, tad turpinot ar otrās rindas un pirmās kolonnas pikseli, utt.

4.5.2. Sistēmas NDVI FS praktiskā realizācija

Sistēmas praktiskā realizācija lietojumprogrammas veidā tiek veikta, izmantojot augsta līmeņa programmēšanas valodu un skaitļošanas vidi *MATLAB*. Sistēma tiek realizēta, izmantojot savstarpēji saistītu funkciju un skripta kopumu biznesa loģikas nodrošināšanai, rakstzīmju orientēto lietotāja saskarni un datu krātuvi. Datu krātuve tiek realizēta datu failu veidā, kas tiek glabāti lietotāja datu nesējā, piemēram, uz datora cieta diska. Skripts tiek izmantots, lai uzsāktu sistēmas darbu un noteiktā secībā izsauktu visas nepieciešamas funkcijas, sākot ar lietotāja parametru ievadīšanu.

Sistēmas biznesa loģika ietver noteiktas *NDVI* laika rindas īstermiņa prognozes aprēķināšanu, izmantojot lietotāja ievadītos datus, datu krātuvē esošos *MODIS NDVI*

attēlus, kā arī noteiktus algoritmus. Biznesa loģikas nodrošināšanai tiek izmantota virkne dažādu *MATLAB* funkciju, gan iebūvēto, gan autora izstrādāto. Modificētās *VMD* metodes *VMD_modified* koda izstrādāšana tiek veikta, par pamatu izmantojot oriģinālās *VMD* metodes *MATLAB* funkciju *VMD*, ko ir izveidojuši šīs metodes autori [119]. Darba autors no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieejas funkcijas kodu izstrādā, izmantojot *MATLAB* iebūvēto fundamentālo atrisinājumu sistēmas aprēķināšanas metodi (funkcija *null*), nelineāro regresijas analīzi ar mazāko kvadrātu metodi (funkcija *lsqcurvefit*) un lineāro regresijas analīzi ar mazāko kvadrātu metodi (funkcija *lsqlin*). Četrkārtīgas precizitātes (*quadruple precision*) sasniegšanai *MATLAB* vidē tiek izmantots uzlabotais multiprecizitātes rīks [2].

Pakāpeniskās regresijas analīzes veikšanai tiek izmantota iebūvētā funkcija *stepwisefit*. Fāžu telpas rekonstrukcija ar laika aiztures metodi tiek realizēta ar funkcijas *phasespace* palīdzību, kas ietilpst haotisko sistēmu rīkā [54]. Pakāpeniskā regresijas analīze tiek veikta, izmantojot *MATLAB* dimensiju skaita reducēšanas rīku [59]. Rekurento slāņu neironu tīkla (*LRNN*) modeļa apmācībai tiek izmantota iebūvētā funkcija *layrecnet*. Visu nepieciešamo datu pirmapstrādes parametru piemēroto vērtību meklēšana, kā arī rekurento slāņu neironu tīkla parametru piemēroto vērtību meklēšana tiek realizēta, izmantojot autora izstrādātas *MATLAB* funkcijas.

4.6. Kopsavilkums un secinājumi

Promocijas darba gaitā ir izstrādāta prognozēšanas sistēma *NDVI FS*, kas veic normalizēta veģetācijas indeksa (*NDVI*) īstermiņa prognozēšanu vienu laika soli uz priekšu. Nodaļā tiek apskatīta šīs sistēmas arhitektūra, darbības princips un praktiskā realizācija.

Izmantojot lietotāja saskarni, lietotājs ievada koordinātas (*X* un *Y*), kā arī no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieejas izvēli (izmantot vai neizmantot). Datu pirmapstrādes moduli veido četri bloki, kur blokā "DP1" tiek veikta *VMD* apakšsignāla aproksimācija, blokā "DP2" tiek rekonstruētas fāžu telpas, datu kopu izveidošanas blokā tiek iegūtas apmācības, validācijas un testa datu kopas, savukārt blokā "DP3" tiek veikta informatīvāko atribūtu atlasīšana un lineāri nekorelējošu atribūtu kopu izveidošana. Mašīnmācīšanās modulis sastāv no apmācības bloka un prognozēšanas bloka. Apmācības blokā tiek iegūts konkrētajai *NDVI* laika

rindai piemērots prognozēšanas modelis. Bet prognozēšanās blokā tiek veikta jaunas *NDVI* laika rindas vērtības prognozēšana vienu soli uz priekšu, kas tiek izsniegta lietotājam. Galvenie secinājumi par izstrādāto prognozēšanas sistēmas arhitektūru un darbības principu ir šādi.

1. Sistēmas izejā esošā *NDVI* prognoze var tikt izmantota par ieeju lēmumu atbalsta sistēmai, ar kuras palīdzību lauksaimnieki var savlaicīgi pieņemt nepieciešamos lēmumus.
2. Sistēma ļauj iegūt augstas precizitātes *NDVI* laika rindas prognozes, ieejā izmantojot tikai *NDVI* vēsturiskos novērojumus.
3. Visa datu pirmapstrāde, *LRNN* modeļa apmācība un laika rindas prognozēšana ir pilnībā automatizēta.

Sistēma praktiski tiek realizēta lietojumprogrammas veidā, izmantojot augsta līmeņa programmēšanas valodu un skaitļošanas vidi *MATLAB*. Sistēma tiek izstrādāta, izmantojot savstarpēji saistītu funkciju un skripta kopumu biznesa loģikas nodrošināšanai, rakstzīmju orientēto lietotāja saskarni un datu krātuvi, kas tiek realizēta datu failu veidā uz lietotāja datu nesēja. Sistēmas biznesa loģikas nodrošināšanai, kas ietver *NDVI* laika rindas īstermiņa prognozes iegūšanu, izmantojot *MODIS NDVI* attēlus un lietotāja ievadītos parametrus, tiek izmantotas gan *MATLAB* iebūvētās funkcijas, gan autora izstrādātās funkcijas.

5. IZSTRĀDĀTĀS PROGNOZĒŠANAS SISTĒMAS PRECIZITĀTES NOVĒRTĒŠANA

Promocijas darbā izstrādātās normalizēta veģetācijas indeksa (*Normalized Difference Vegetation Index, NDVI*) laika rindu prognozēšanas sistēmas *NDVI FS* precizitātes novērtēšanai nepieciešams pārbaudīt, kādi prognozēšanas rezultāti tiks iegūti, lietojot izstrādāto rīku, un salīdzināt tos ar klasiskajām prognozēšanas metodēm.

5.1. Eksperimentu plāns un apraksts

Prognozēšanas sistēmas *NDVI FS* precizitātes novērtēšanai ir izstrādāts eksperimentu plāns, kurā ir iekļauti septiņi eksperimenti. Pirmo piecu eksperimentu mērķis un apraksts ir dots 5.1. tabulā. Šie eksperimenti ir nepieciešami, lai novērtētu un salīdzinātu prognozēšanas precizitāti, kas ir iegūta, izmantojot izstrādāto sistēmu *NDVI FS* un izmantojot klasiskās prognozēšanas metodes. Šajos piecos eksperimentos tiek izmantotas 100 *NDVI* laika rindas, kur katrai laika rindai prognozēšana tiek veikta atsevišķi. Laika rindas tiek izvēlētas, ģenerējot gadījuma skaitļus – pikseļu rindas un kolonnas numurus vidējās izšķirtspējas attēlveidošanas (*Moderate Resolution Imaging Spectroradiometer, MODIS*) *NDVI* attēlos. Nejauši izvēlētas 100 *NDVI* laika rindas visos šajos piecos eksperimentos ir vienas un tās pašas, lai objektīvi salīdzinātu iegūto prognozēšanas precizitāti. Precizitātes novērtēšana tiek veikta, aprēķinot trīs dažādu zaudējumu funkciju: saknes vidējās kvadrātiskās kļūdas (*Root Mean Square Error, RMSE*) virzienu simetrijas (*Directional Symmetry,* DS) un pielāgota determinācijas koeficienta R^2_{adj} vērtības.

Pirmajā, otrajā un trešajā eksperimentā tiek novērtēta klasisko prognozēšanas metožu: vienkāršā slīdošā vidējā (*Simple Moving Average, SMA*), varbūtiskās metodes – diskrētā laika, nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēžu un lineārās metodes – automātiski regresējoša integrētā slīdošā vidējā (*Autoregressive Integrated Moving Average, ARIMA*) precizitāte *NDVI* laika rindu prognozēšanas uzdevumā. Ceturtajā un piektajā eksperimentā tiek novērtēta izstrādātā sistēma *NDVI* laika rindu prognozēšanas uzdevumā, izmantojot un neizmantojot autora izstrādāto apakšsignāla aproksimācijas pieeju. Ceturtais un piektais eksperiments ir nepieciešami, lai apstiprinātu vai noliegtu promocijas darbā izvirzīto pirmo hipotēzi.

5.1. tabula

Eksperimentu plāns *NDVI FS* precizitātes salīdzināšanai ar citām metodēm

Nr.	Eksperimenta apraksts	Mērķis
1.	Fāžu telpas rekonstrukcija tiek izmantota, lai sadalītu *NDVI* laika rindu fragmentos. No katra fragmenta tiek aprēķināta vidējā vērtība, kas ir prognoze vienu laika soli uz priekšu.	Izpētīt vienkāršā slīdošā vidējā (*SMA*) precizitāti *NDVI* laika rindu prognozēšanas uzdevumā.
2.	Fāžu telpas rekonstrukcija tiek izmantota, lai sadalītu *NDVI* laika rindu fragmentos. No katra fragmenta tiek atlasīti informatīvākie atribūti. Katram fragmentam tiek atrasti līdzīgākie fragmenti, un tiek iegūta katra fragmenta nākamā vērtība. Prognoze tiek aprēķināta kā visu šo nākamo vērtību vidējā vērtība.	Izpētīt nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēžu precizitāti *NDVI* laika rindu prognozēšanas uzdevumā.
3.	No *NDVI* laika rindas vēsturiskajiem novērojumiem secīgi tiek atlasīti katri 70 % novērojumu, uz kuriem apmācās *ARIMA* modelis, un tad tiek prognozēta viena laika rindas vērtība uz priekšu.	Izpētīt automātiski regresējoša integrētā slīdošā vidējā (*ARIMA*) metodes precizitāti *NDVI* laika rindu prognozēšanas uzdevumā.
4.	Fāžu telpas rekonstrukcija tiek izmantota, lai sadalītu *NDVI* laika rindu fragmentos. No katra fragmenta tiek atlasīti informatīvākie atribūti, un tiek veikta katra fragmenta transformācija, lai iegūtu lineāri nekorelējošus atribūtus. Prognoze tiek iegūta, izmantojot *LRNN* modeli, kam ieejā tiek padoti pirmapstrādāti fragmenti.	Izpētīt *NDVI* prognozēšanas sistēmas precizitāti *NDVI* laika rindu prognozēšanas uzdevumā, neizmantojot no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieeju.
5.	Izmantojot apakšsignāla aproksimācijas pieeju, iegūst tuvinātu apakšsignālu. Fāžu telpas rekonstrukcija tiek izmantota, lai sadalītu *NDVI* laika rindu un apakšsignālu fragmentos. No katra fragmenta tiek atlasīti informatīvākie atribūti, un tiek veikta katra fragmenta transformācija, lai iegūtu lineāri nekorelējošus atribūtus. Prognoze tiek iegūta, izmantojot *LRNN* modeli, kam ieejā tiek padoti pirmapstrādāti fragmenti.	Izpētīt *NDVI* prognozēšanas sistēmas precizitāti *NDVI* laika rindu prognozēšanas uzdevumā, izmantojot no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieeju.

Sestā un septītā eksperimenta mērķis un apraksts ir dots 5.2. tabulā. Sestajā eksperimentā tiek novērtēta iespēja izmantot izstrādātajā sistēmā iegūtos pirmapstrādes parametrus un apmācīto prognozēšanas modeli, lai prognozētu citas *NDVI* laika rindas, kur visi pirmapstrādes parametri un prognozēšanas modelis tiek iegūti no izstrādātās prognozēšanas sistēmas *NDVI FS*, kas neizmanto aproksimācijas pieeju. Septītajā eksperimentā tiek novērtēta iespēja izmantot datu pirmapstrādes parametrus un prognozēšanas modeli, lai prognozētu citas *NDVI* laika rindas, kur visi pirmapstrādes parametri un prognozēšanas modelis tiek iegūti no izstrādātās prognozēšanas sistēmas *NDVI FS*, kas izmanto apakšsignāla aproksimācijas pieeju.

5.2. tabula

Eksperimentu plāns pirmapstrādes parametru un prognozēšanas modeļu pārnešanai

Nr.	**Eksperimenta apraksts**	**Mērķis**
6.	Eksperimentā tiek izmantotas *NDVI* laika rindas, kas ir iegūtas no 11x11 pikseļu režģa. Uz centrālā pikseļa laika rindas tiek veikta apmācība. Izstrādātajā prognozēšanas sistēmā *NDVI FS*, kas neizmanto aproksimācijas pieeju, tiek iegūti pirmapstrādes parametri un apmācītais prognozēšanas modelis. Tie tiek izmantoti, lai prognozētu pārejās 120 *NDVI* laika rindas bez jaunas datu pirmapstrādes un apmācības.	Izpētīt prognozēšanas modeļu pārnešanu uz kaimiņu pikseļu *NDVI* laika rindām, veicot to prognozēšanu bez jaunas apmācības un neizmantojot apakšsignāla aproksimācijas pieeju.
7.	Eksperimentā tiek izmantota 121 *NDVI* laika rinda, kas iegūtas no 11x11 pikseļu režģa. Centrālā pikseļa laika rinda tiek izmantota apmācībai. Izstrādātajā prognozēšanas sistēmā *NDVI FS*, kas izmanto aproksimācijas pieeju, tiek iegūti pirmapstrādes parametri un apmācītais prognozēšanas modelis. Tie tiek izmantoti, lai prognozētu pārejās 120 *NDVI* laika rindas bez jaunas datu pirmapstrādes un apmācības.	Izpētīt prognozēšanas modeļu pārnešanu uz kaimiņu pikseļu *NDVI* laika rindām, veicot to prognozēšanu bez jaunas apmācības un izmantojot apakšsignāla aproksimācijas, pieeju.

Sestais un septītais eksperiments ir nepieciešams, lai apstiprinātu vai noliegtu promocijas darbā izvirzīto otro hipotēzi.

5.2. *NDVI* prognozēšana ar vienkāršā slīdošā vidējā izmantošanu

Pirmā eksperimenta mērķis ir izpētīt vienkāršā slīdošā vidējā precizitāti *NDVI* laika rindu prognozēšanas uzdevumā. Eksperimentā *NDVI* laika rindai pēc vienādojuma (2.4.) tiek rekonstruēta fāžu telpa ar dažādiem dimensijas m lielumiem. Fāžu telpas dimensija m tiek meklēta intervālā [1; 50], bet laika aizture $\tau = 1$. Katram rekonstruētās fāžu telpas aiztures vektoram tiek aprēķināta vidējā vērtība, kas arī ir prognoze vienam laika solim uz priekšu, tad tiek aprēķinātas zaudējumu funkciju *RMSE* (izteiksme (1.3.)), *DS* (izteiksme (1.4.)) un pielāgotā R^2_{adj} (izteiksme (1.6.)) vērtības. Par piemēroto dimensiju m, un attiecīgi vienkāršā slīdošā vidējā (*SMA*) periodu (skat. 2.6.1. punktu) tiek uzskatīta tā vērtība, pie kuras tiek sasniegta minimālā zaudējumu funkcijas *RMSE* vērtība. Eksperiments tiek atkārtots 100 reižu, katru reizi ņemot citu *NDVI* laika rindu, un šo triju zaudējumu funkciju: *RMSE*, *DS* un pielāgotā R^2_{adj} vērtību vidējās, minimālās un maksimālās vērtības 100 gadījumiem dotas 5.3. tabulā. Savukārt visas iegūtās zaudējumu funkcijas *RMSE* vērtības 100 eksperimenta gadījumiem ir dotas 3. pielikumā.

5.3. tabula

Ar *SMA* iegūtas zaudējumu funkciju vērtības

Vērtība	***RMSE***	***DS***	R^2_{adj}
Vidējā vērtība	0,0442	93,40 %	0,94
Minimālā vērtība	0,0305	89,27 %	0,90
Maksimālā vērtība	0,0605	95,78 %	0,98

Attiecīgi 5.3. tabulā pirmā kolonna atbilst *RMSE* vērtību vidējai vērtībai 100 eksperimenta gadījumos, kur katra *RMSE* vērtība parāda saknes vidējo kvadrātisko kļūdu starp katru novēroto un prognozēto *NDVI* laika rindu, kā arī tiek dotas minimālā un maksimālā *RMSE* vērtība. Savukārt otrajā un trešajā kolonnā līdzīgi tiek aprakstīta statistika par zaudējumu funkcijām *DS* un pielāgoto determinācijas koeficientu R^2_{adj}.

5.3. Eksperimenta veikšana ar Markova ķēžu izmantošanu

Otrā eksperimenta mērķis ir izpētīt nepārtrauktu stāvokļu kopas Markova ķēžu precizitāti *NDVI* laika rindu prognozēšanas uzdevumā. Darba autors eksperimentu ar diskrētā laika, diskrētās stāvokļu kopas Markova ķēžu izmantošanu *NDVI* laika rindu prognozēšanā ir aprakstījis pētījumā [91], bet *m-tās* kārtas diskrētā laika, nepārtrauktu stāvokļu kopas Markova ķēžu izmantošana *NDVI* laika rindu prognozēšanā ir aprakstīta darba autora pētījumā [93].

Eksperimentā tiek izmantota nepārtrauktu stāvokļu kopas Markova ķēdes (skat. 2.6.2. punktu). Eksperimenta gaitā no katras *NDVI* laika rindas tiek iegūta fāžu telpa ar dimensiju 50, kura ir sākotnējā ieejas datu kopa. Ieejas datu kopa tiek sadalīta atribūtu kopā un prognozējamajā parametrā. Gan atribūtu kopa, gan prognozējamais parametrs tiek sadalīti trīs daļās, iegūstot apmācības, validācijas un testa datu kopas, kur 70 % no datu ierakstiem tiek izmantoti kā apmācības datu kopa, 15 % tiek izmantoti kā validācijas datu kopa un pēdējie 15 % no datu ierakstiem tiek izmantoti kā testa datu kopa.

Lai samazinātu ieejas datu dimensiju skaitu un uzlabotu nepārtraukta stāvokļu kopas Markova ķēdes prognozējamību, pakāpeniskā regresijas analīze tiek pielietota atribūtu atlasīšanai. Atribūta pievienošanai datu kopai tiek izmantota p vērtība 0,05, tāpat arī atribūta izņemšanai no datu kopas tiek izmantota p vērtība 0,05. Informatīvāko atribūtu indeksi tiek izmantoti, lai atlasītu atribūtu apakškopas no apmācības, validācijas un testa atribūtu kopām.

Pēc datu pirmapstrādes, nepārtrauktu stāvokļu kopas *m-tās* kārtas (kur m ir atribūtu skaits datu kopās pēc datu pirmapstrādes veikšanas) Markova ķēde tiek izmantota, lai iegūtu varbūtiskās prognozes. Lai atrastu vektorus no apmācības atribūtu kopas, kur Eiklīda distance starp pēdējo vektoru un jebkuru no iepriekšējiem vektoriem ir mazāka vai vienāda, nekā puse no diametra ε, tiek izmantota šķērsvalidācija. Diametrs ε tiek meklēts intervālā [0,1; 0,6].

Katram vektoram, kurš pieder atlasītajai vektoru kopai, tiek atrasts nākamais laika rindas novērojums, kas sākotnējā fāžu telpā seko atbilstošajam laika rindas fragmentam. Šiem laika rindas novērojumiem tiek aprēķināta vidējā vērtība, un tā ir iegūtā prognoze. Tālāk eksperiments tiek atkārtots ar pārējām *NDVI* laika rindām.

Kā precizitātes kritēriji tiek izvēlētas trīs zaudējumu funkcijas: saknes vidējā kvadrātiskā kļūda (*RMSE*), virzienu simetrija (*DS*) un pielāgotais determinācijas

koeficients R^2_{adj}. Eksperimentos aprēķinātās prognožu zaudējumu funkciju: *RMSE*, *DS* un pielāgotā R^2_{adj} vidējās, minimālās un maksimālās vērtības dotas 5.4. tabulā. Aprēķinātās zaudējumu funkcijas *RMSE* vērtības 100 eksperimenta gadījumiem (100 *NDVI* laika rindām) ir dotas 3. pielikumā.

5.4. tabula

Ar Markova ķēdēm iegūtās zaudējumu funkciju vērtības

Datu kopa	**Vērtības**	***RMSE***	***DS***	R^2_{adj}
Validācijas datu kopa	Vidējā vērtība	0,0210	86,70 %	0,96
	Minimālā vērtība	0,0116	70,96 %	0,91
	Maksimālā vērtība	0,0345	100,00 %	0,99
Testa datu kopa	Vidējā vērtība	0,0214	83,76 %	0,90
	Minimālā vērtība	0,0084	56,66 %	0,63
	Maksimālā vērtība	0,0813	100,00 %	0,98

Iegūtie rezultāti 5.4. tabulā parāda, ka ir sasniegta labāka prognozēšanas precizitāte, salīdzinot ar vienkāršā slīdošā vidējā sasniegtajiem rezultātiem.

5.4. *NDVI* prognozēšana ar *ARIMA*

Trešā eksperimenta mērķis ir izpētīt *ARIMA* metodes precizitāti *NDVI* laika rindu prognozēšanas uzdevumā. Eksperimentā *NDVI* laika rindai atlasa pēdējos 30 % no laika rindas novērojumiem, un izveido testa datu kopu. Eksperimenta gaitā autoregresīvā polinoma kārtā p un slīdošā vidējā polinoma kārtā q (skat. 2.6.3. punktu) tiek meklētas intervālā [0; 5]. Pie katras no p un q kombinācijām tiek atlasīti pirmie 70 % no *NDVI* laika rindas novērojumiem. Dikija-Fullera tests tiek veikts, lai pārbaudītu, vai laika rinda, kas sastāv no apmācības datu kopas novērojumiem, ir stacionāra. Kamēr tests nespēj noraidīt nulles hipotēzi par vienības saknes eksistēšanu, laika rinda tiek diferencēta. Kad ir atrasta d vērtība, kas ļauj iegūt stacionāru diferenču laika rindu, tiek izveidots $ARIMA(p,d,q)$ modelis, un tiek veikta prognoze vienu laika rindas vienību uz priekšu. Tad tiek atlasīti nākamie 70 % no *NDVI* laika rindas novērojumiem, sākot no laika soļa $t=2$, un tiek veikta prognozēšana nākamajam laika rindas elementam. Tad q vērtība palielinās par viens, un process atkārtojas. Kad q

sasniedz vērtību pieci, tad tālāk p vērtība palielinās par viens, un q pieņem vērtību nulle. Beigās par piemērotākajām tiek izvēlētas tās p un q vērtības, pie kurām zaudējumu funkcijas *RMSE* vērtība starp prognozētajām un novērotajām testa datu kopas vērtībām ir minimāla.

Kad ir prognozēti visi testa datu kopas novērojumi, tad eksperiments tiek atkārtots nākamajai *NDVI* laika rindai. Kopā eksperiments tiek atkārtots 100 reizes. Kā precizitātes kritērijs tiek izvēlētas trīs zaudējumu funkcijas: saknes vidējā kvadrātiskā kļūda (*RMSE*), virzienu simetrija (*DS*) un pielāgotais determinācijas koeficients R^2_{adj}. Šajos 100 eksperimenta gadījumos iegūtās zaudējumu funkcijas: *RMSE*, *DS* un R^2_{adj} vidējās, minimālās un maksimālās vērtības uz testa datu kopas dotas 5.5. tabulā. Zaudējumu funkcijas *RMSE* vērtības 100 eksperimenta gadījumiem tiek parādītas 3. pielikumā.

5.5. tabula

Zaudējumu funkciju vidējas vērtības 100 *NDVI* laika rindu prognozēm

Vērtības	***RMSE***	***DS***	R^2_{adj}
Vidējā vērtība	0,0108	93,85 %	0,96
Minimālā vērtība	0,0004	85,89 %	0,93
Maksimālā vērtība	0,0345	98,71 %	0,98

Kā redzams 5.5. tabulā, iegūtā prognozēšanas precizitāte ir labāka, salīdzinot gan ar vienkāršā slīdošā vidējā sasniegtajiem rezultātiem, gan ar diskrētā laika, nepārtrauktu stāvokļu kopas *m*-tās kārtas Markova ķēžu sasniegtajiem rezultātiem.

5.5. Eksperimentu ar sistēmu *NDVI FS* raksturojums

Darba autors eksperimentus ar Elmana rekurento neironu tīklu izmantošanu *NDVI* laika rindu prognozēšanai ir aprakstījis pētījumā [92], eksperimenti ar *LRNN* izmantošanu *NDVI* laika rindu prognozēšanai ir aprakstīti autora pētījumā [87], bet prognozēšanas sistēmas, kura neizmanto uz *VMD* balstīto aproksimācijas pieeju, izmantošana *NDVI* laika rindu prognozēšanai ir aprakstīta autora pētījumos [88], [94].

5.5.1. Prognozēšana bez VMD bāzētās aproksimācijas pieejas

Normalizētā veģetācijas indeksa (*NDVI*) laika rindai tiek iegūta fāžu telpa ar dimensiju $m = 50$ un laika aizturi $\tau = 1$, un tādējādi tiek iegūta sākotnējā ieejas datu kopa. Ieejas datu kopa tiek iedalīta divās daļās: atribūtu kopā un prognozējamajā parametrā. Gan atribūtu kopa, gan prognozējamais parametrs tiek iedalīti trīs daļās, iegūstot apmācības, validācijas un testa datu kopas, kur 70 % no datu ierakstiem tiek izmantoti kā apmācības datu kopa, 15 % no ierakstiem tiek izmantoti kā validācijas datu kopa un pēdējie 15 % no ierakstiem tiek izmantoti kā testa datu kopa.

Lai samazinātu ieejas datu dimensiju skaitu, pakāpeniskā regresijas analīze tiek pielietota atribūtu atlasīšanai apmācības datu kopā. Eksperimentā tiek atrasta tāda atribūtu apakškopa, kas minimizē *RMSE* vērtību, prognozējot apmācības datu kopas prognozējamo parametru. Atribūta pievienošanai datu kopai tiek izmantota *p* vērtība 0,05; un tāda pat *p* vērtība tiek izmantota atribūta izņemšanai no datu kopas. Šādas atribūtu apakškopas, izmantojot atrastos informatīvāko atribūtu indeksus, tiek atlasītas arī no validācijas un testa atribūtu kopām. Principiālo komponentu analīze tiek izmantota, lai iegūtu lineāri nekorelējošu atribūtu kopas.

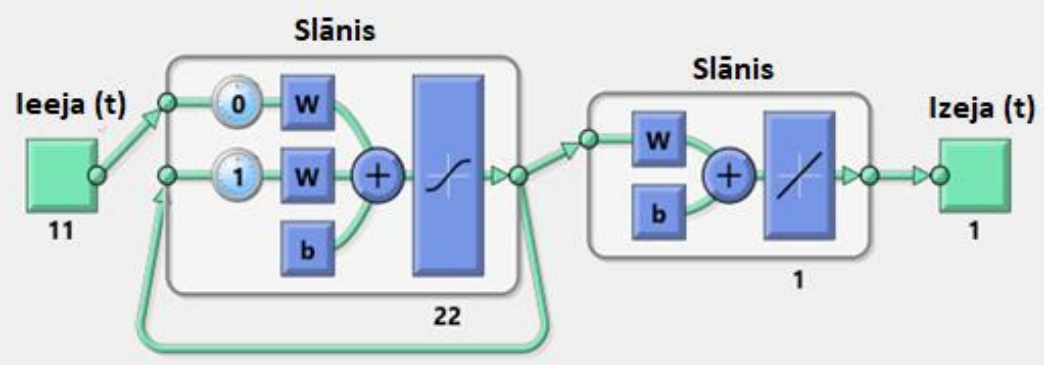
Rekurento slāņu neironu tīklu (*LRNN*) modeļi eksperimentā tiek apmācīti ar Levenberga-Markvardta kļūdas atgriezeniskās izplatības algoritmu ar Beijesa regularizāciju, kas ir ātrs apmācības algoritms un ļauj iegūt neironu tīkla modeli ar labām vispārināšanas spējām. Neironu tīklu svari un konstantes tiek inicializētas ar nelieliem gadījuma skaitļiem intervālā [–0,1; 0,1]. Slēpto slāņu skaits ir viens visiem neironu tīklu modeļiem. Hiperboliskā tangensa funkcija un lineārā funkcija tiek izmantotas kā aktivizācijas funkcijas atbilstoši slēptā un izejas slāņa neironiem. Epohu skaits, kas tiek izmantotas *LRNN* apmācībai, ir 10000.

Tā kā neironu skaits slēptajā slānī ir svarīgs parametrs, kas ietekmē prognozēšanas precizitāti, tad, izmantojot šķērsvalidāciju, tiek atrasts katram *LRNN* modelim piemērots slēptā slāņa neironu skaits, kas minimizē *RMSE* vērtību validācijas datu kopā. Neironu skaits slēptajā slānī tiek meklēts intervālā [1; 50].

Lai uzlabotu neironu tīkla vispārināšanas spējas, tiek izmantota agrīnās apstāšanās metode. Kad *LRNN* sāk pārmērīgi pielāgoties datiem (*overfitting*), globālā kļūda validācijas datu kopā sāk pieaugt. Kad kļūda validācijas datu kopā pieaug 100 epohas pēc kārtas, apmācība tiek apturēta, un tiek paņemtas tās svaru un

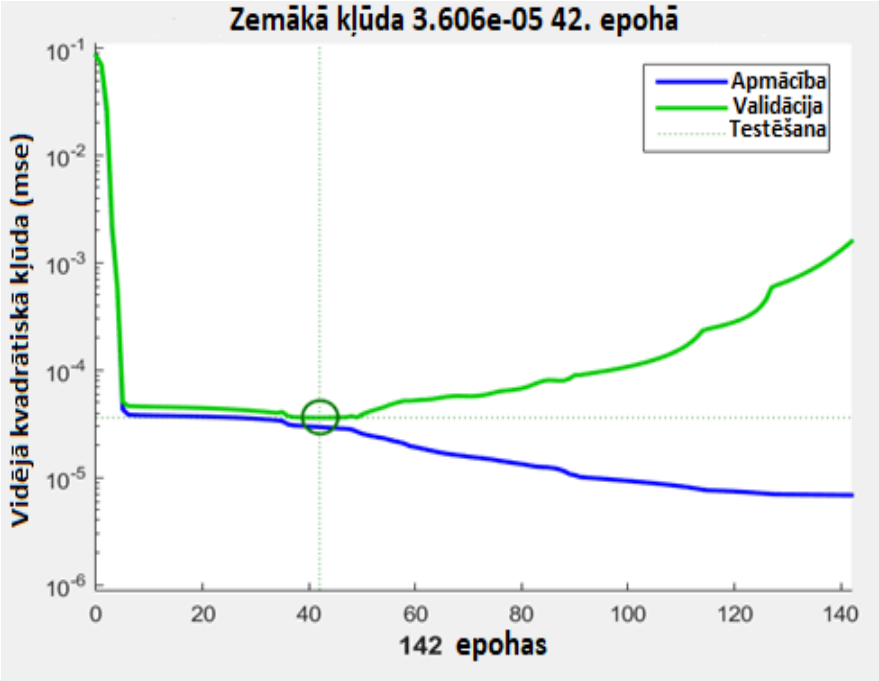
konstanšu vērtības, kas tika iegūtas pie minimālas kļūdu vērtības 101 epohu atpakaļ. Eksperiments tiek atkārtots visām 100 laika rindām.

Rekurento slāņu neironu tīkla piemērotākā modeļa topoloģija vienai no eksperimentā izmantotajām *NDVI* laika rindām dota 5.1. attēlā.



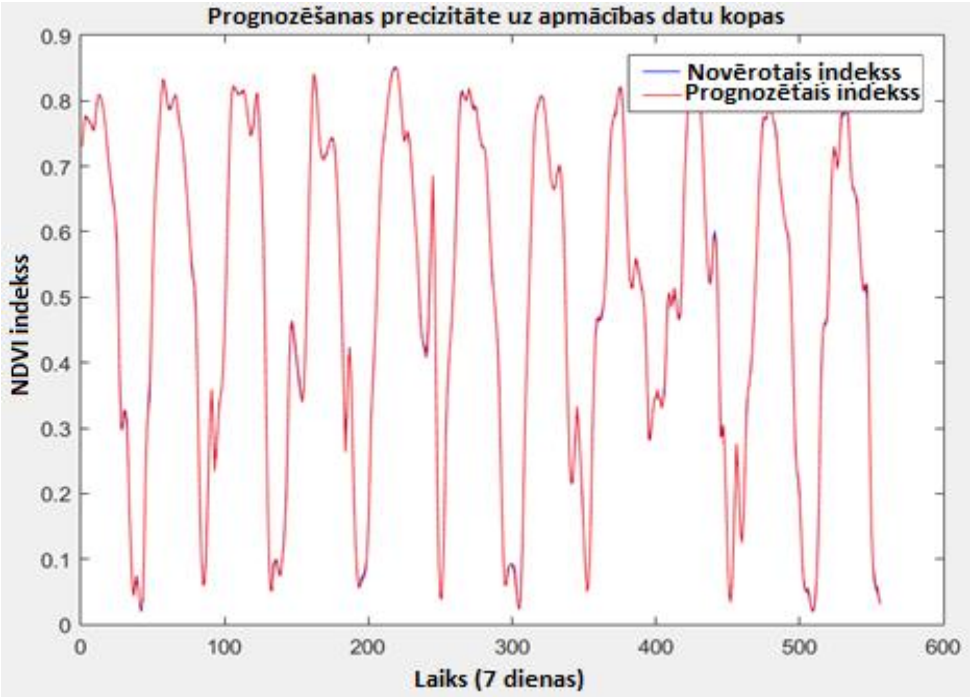
5.1. att. Piemērotākā *LRNN* topoloģija vienai no *NDVI* laika rindām.

Konverģence labākajam *LRNN* modelim vienai no *NDVI* laika rindām, kas sasniegta pēc 42 epohu apmācības parādīta 5.2. attēlā.



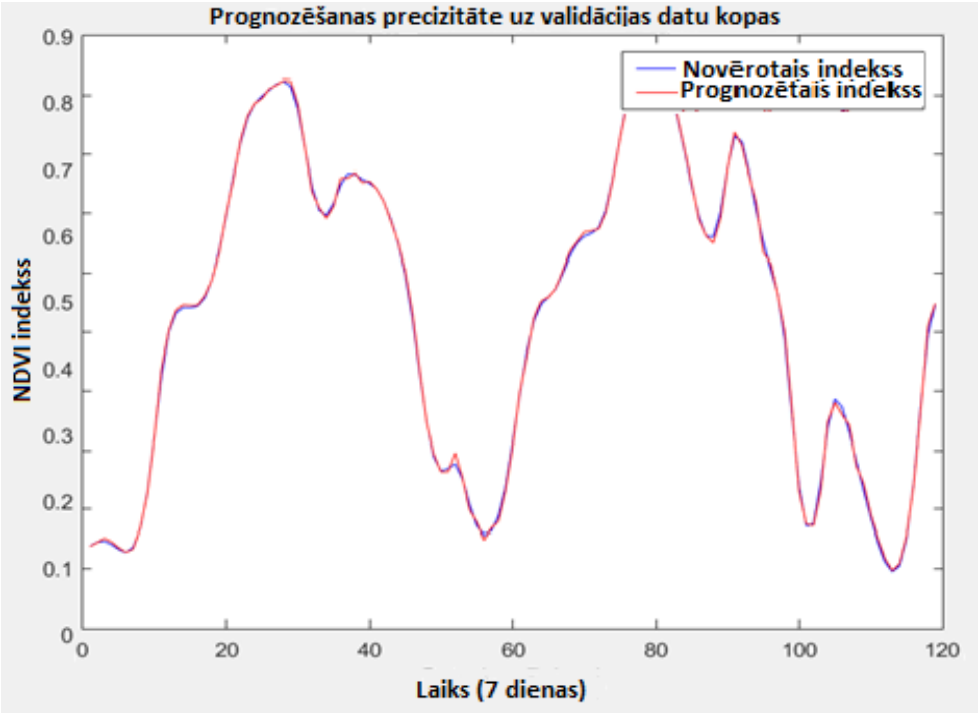
5.2. att. Labākā *LRNN* modeļa konverģence vienai no *NDVI* laika rindām.

Vienas no eksperimentā izmantotās *NDVI* laika rindas novērotās un prognozētās vērtības apmācības datu kopā parādītas 5.3. attēlā. Katra *NDVI* vērtība 5.3. attēlā tiek iegūta reizi katrās septiņās dienās.



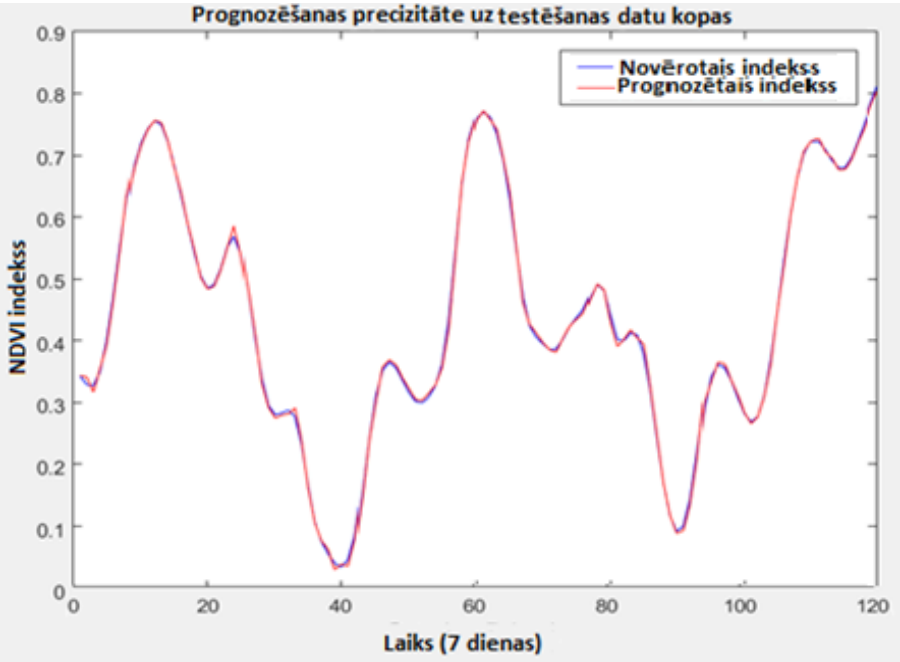
5.3. att. Novērotās un prognozētās *NDVI* vērtības apmācības datu kopā.

Novērotās un prognozētās *NDVI* laika rindas vērtības validācijas datu kopā parādītas 5.4. attēlā.



5.4. att. Novērotās un prognozētās *NDVI* vērtības validācijas datu kopā.

Novērotās un prognozētās *NDVI* laika rindas vērtības testēšanas datu kopā parādītas 5.5. attēlā.



5.5. att. Novērotās un prognozētās *NDVI* vērtības testa datu kopā.

Eksperimentā ar *NDVI FS* aprēķinātās zaudējumu funkciju: *RMSE*, *DS* un pielāgotā R^2_{adj} vidējās, minimālās un maksimālās vērtības bez autora piedāvātās aproksimācijas pieejas izmantošanas dotas 5.6. tabulā. Visos 100 eksperimenta gadījumos aprēķinātās zaudējumu funkcijas *RMSE* vērtības ir dotas 3. pielikumā.

5.6. tabula

Prognožu zaudējumu funkciju vērtības 100 *NDVI* laika rindām

Datu kopa	Vērtības	*RMSE*	*DS*	R^2_{adj}
Apmācības datu kopa	Vidējā vērtība	0,0015	97,52 %	0,99
	Minimālā vērtība	0,0009	91,21 %	0,99
	Maksimālā vērtība	0,0037	100,00 %	0,99
Validācijas datu kopa	Vidējā vērtība	0,0017	97,43 %	0,99
	Minimālā vērtība	0,0007	90,00 %	0,99
	Maksimālā vērtība	0,0045	100,00 %	0,99
Testa datu kopa	Vidējā vērtība	0,0012	97,61 %	0,99
	Minimālā vērtība	0,0004	84,37 %	0,99
	Maksimālā vērtība	0,0038	100,00 %	0,99

Iegūtās zaudējumu funkciju vērtības 5.6. tabulā uzrāda labāku prognozēšanas precizitāti, salīdzinot ar vienkāršā slīdošā vidējā sasniegtajiem rezultātiem, ar diskrētā laika, nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēžu sasniegtajiem rezultātiem un ar *ARIMA* modeli sasniegtajiem rezultātiem.

5.5.2. Prognozēšana ar VMD apakšsignāla aproksimācijas pieeju

Normalizētā veģetācijas indeksa (*NDVI*) laika rindai tiek iegūts atbilstošs apakšsignāls, izmantojot darba autora piedāvāto variāciju apakšsignālu dekompozīcijas (*VMD*) metodes modifikāciju kopā ar no šīs metodes iegūtā apakšsignāla vērtību aproksimācijas pieeju, kura ir aprakstīta 3. nodaļā.

Modificētā *VMD* metode izmanto šādus parametrus: laika solis $\tau = 0{,}000001$, frekvenču inicializācijas parametrs $DC = 1$, konverģences kritērijs $\varepsilon = 0$, apakšsignālu skaits $K = 1$. Piemērots apakšsignālu summas precizitāti regulējošs parametrs α tiek meklēts intervālā [1; 10] ar soli 0,01. Lai atrastu piemērotu parametru α, no modificētās *VMD* metodes iegūtajam apakšsignālam tiek rekonstruēta fāžu telpa ar $m = 50$ un laika aizturi $\tau = 1$, tāpat arī *NDVI* laika rindai tiek rekonstruēta fāžu telpa ar šiem pašiem parametriem. Fāžu telpas tiek apvienotas, iegūstot ieejas datu kopu, kura tiek sadalīta divās daļās: atribūtu kopā un prognozējamajā parametrā. Gan atribūtu kopa, gan prognozējamais parametrs tiek sadalīti trīs daļās, iegūstot apmācības, validācijas un testa datu kopās, kur 70 % no datu ierakstiem tiek izmantoti kā apmācības datu kopa, 15 % no ierakstiem tiek izmantoti kā validācijas datu kopa un pēdējie 15 % no ierakstiem tiek izmantoti kā testa datu kopa. Apmācības datu kopa tiek padota ieejā pakāpeniskās regresijas modelim, kurš prognozē *NDVI* vērtības, atrodot minimālu *RMSE* vērtību, atlasot iespējami mazāku atribūtu apakškopu. Katrai iespējamai parametra α vērtībai tiek iegūta minimālā *RMSE* vērtība uz testa datu kopas, ko ļauj sasniegt apmācītais pakāpeniskās regresijas modelis, un savukārt minimālā vērtība no visām šīm minimālajām *RMSE* vērtībām arī ļauj iegūt piemēroto apakšsignālu summas precizitāti regulējošā parametra α vērtību. Parametra α meklēšanai netiek izmantots rekurento slāņu neironu tīkls (*LRNN*), jo tas prasa ilgu apmācības laiku, jo parametram α ir iespējama 901 vērtība. Bez parametra α piemērotās vērtības meklēšanas būtu jānosaka arī citu parametru piemērotās vērtības *LRNN* modelī. Rezultātā *LRNN* piemērota modeļa apmācība būtu laikietilpīga.

Kad piemērotā parametra α vērtība ir atrasta, no modificētās *VMD* metodes tiek iegūts atbilstošais apakšsignāls. Gan oriģinālajai *NDVI* laika rindai, gan iegūtajam apakšsignālam tiek veikta fāžu telpas rekonstrukcija ar dimensiju $m = 50$ un laika aizturi $\tau = 1$. Abas fāžu telpas tiek apvienotas, iegūstot ieejas datu kopu. Ieejas datu kopa tiek iedalīta divās daļās: atribūtu kopā un prognozējamajā parametrā. Gan atribūtu kopa, gan prognozējamais parametrs tiek iedalīti trīs daļās, iegūstot apmācības, validācijas un testa datu kopās, kur 70 % no datu ierakstiem tiek izmantoti kā apmācības datu kopa, 15 % no ierakstiem tiek izmantoti kā validācijas datu kopa un pēdējie 15 % no ierakstiem tiek izmantoti kā testa datu kopa.

Pakāpeniskā regresijas analīze tiek pielietota atribūtu atlasīšanai apmācības datu kopas atribūtu kopā. Eksperimentā tiek atrasta tāda atribūtu apakškopa, kas minimizē *RMSE* kļūdu uz apmācības datu kopas. Atribūta pievienošanai datu kopai tiek izmantota p vērtība 0,05, tāpat arī atribūta izņemšanai no datu kopas tiek izmantota p vērtība 0,05. Principiālo komponentu analīze tiek izmantota, lai iegūtu lineāri nekorelējošu atribūtu kopas.

Rekurento slāņu neironu tīkla modeļi eksperimentā tiek apmācīti ar Levenberga-Markvardta kļūdas atgriezeniskās izplatības algoritmu ar Beijesa regularizāciju. Neironu tīklu svari un konstantes tiek inicializētas ar nelieliem gadījuma skaitļiem intervālā [–0,1; 0,1]. Visiem *LRNN* modeļiem ir viens slēptais slānis. Slēptā un izejas slāņa neironiem tiek izmantotas hiperboliskā tangensa un lineārā aktivizācijas funkcija. Neironu tīklu modeļu apmācībai tiek izmantota 10000 epohas.

Izmantojot šķērsvalidāciju, tiek atrasts katram *LRNN* modelim piemērots slēptā slāņa neironu skaits, kas minimizē *RMSE* kļūdu validācijas datu kopā. Neironu skaits slēptajā slānī tiek meklēts intervālā [1; 50].

Lai uzlabotu neironu tīkla vispārināšanas spējas, tiek izmantota agrīnās apstāšanās metode. Kad globālā kļūda validācijas datu kopā pieaug 100 epohas pēc kārtas, apmācība tiek apturēta, un tiek paņemtas tās svaru un konstanšu vērtības, kas tika iegūtas pie minimālās kļūdas vērtības, kas tika sasniegta 101 epohu atpakaļ. Eksperiments tiek atkārtots 100 reižu ar katru *NDVI* laika rindu. *NDVI* laika rindu prognozēšanas eksperimentā aprēķinātās prognožu zaudējumu funkciju: *RMSE, DS* un pielāgotā R^2_{adj} vidējās, minimālās un maksimālās vērtības ar autora piedāvātās aproksimācijas pieejas izmantošanu dotas 5.7. tabulā. Savukārt visas zaudējumu funkcijas *RMSE* vērtības, kas iegūtas 100 eksperimenta gadījumos, dotas 3. pielikumā.

5.7. tabula

Prognožu zaudējumu funkciju vērtības 100 *NDVI* laika rindām

Datu kopa	Vērtības	*RMSE*	*DS*	R^2_{adj}
Apmācības datu kopa	Vidējā vērtība	0,0008	98,92 %	0,99
	Minimālā vērtība	0,0006	92,85 %	0,99
	Maksimālā vērtība	0,0025	100,00 %	0,99
Validācijas datu kopa	Vidējā vērtība	0,0006	98,33 %	0,99
	Minimālā vērtība	0,0004	96,42 %	0,99
	Maksimālā vērtība	0,0014	100,00 %	0,99
Testa datu kopa	Vidējā vērtība	0,0009	98,80 %	0,99
	Minimālā vērtība	0,0005	96,42 %	0,99
	Maksimālā vērtība	0,0024	100,00 %	0,99

Kā redzams 5.7. tabulā, ir sasniegta labākā prognozēšanas precizitāte pēc visām trim zaudējumu funkciju vērtībām visos piecos eksperimentos.

5.5.3. Prognozēšana ar modeļu pārnešanu bez aproksimācijas pieejas

Sestā eksperimenta ietvaros *NDVI* laika rindu prognozēšana tiek veikta, izmantojot izstrādāto sistēmu *NDVI FS* bez *VMD* apakšsignāla aproksimācijas pieejas izmantošanas, kā tas ir aprakstīts 5.5.2. punktā.

Kad uz nejauši izvēlētas *NDVI* laika rindas ir veikta apmācība, tiek saglabāti visi nepieciešamie datu pirmapstrādes parametri, un tiek iegūts piemērots prognozēšanas modelis, kas uz apmācībā izmantotās *NDVI* laika rindas validācijas datu kopā sniedz mazāko *RMSE* vērtību.

Ap šai *NDVI* laika rindai atbilstošo pikseli satelītattēlā tiek novilkts režģis, kas atbilst noteiktam rādiusam r. Eksperimentā tiek izmantots rādiuss $r = 5$. Tādējādi apkārt izvēlētās *NDVI* laika rindas atbilstošajam pikselim tiek novilkts 11x11 pikseļu liels režģis, kurā bez centrālā pikseļa, uz kura tika veikta apmācība, vēl ir 120 pikseļi. Katram no šiem kaimiņu pikseļiem tiek iegūtas *NDVI* laika rindas.

Vispirms katrai no šīm 120 laika rindām tika veikta fāžu telpas rekonstrukcija ar laika aiztures metodi, tiek veikta atribūtu atlasīšana, izmantojot pakāpenisko regresijas analīzi, un tiek iegūta lineāri nekorelējošu atribūtu kopa, izmantojot

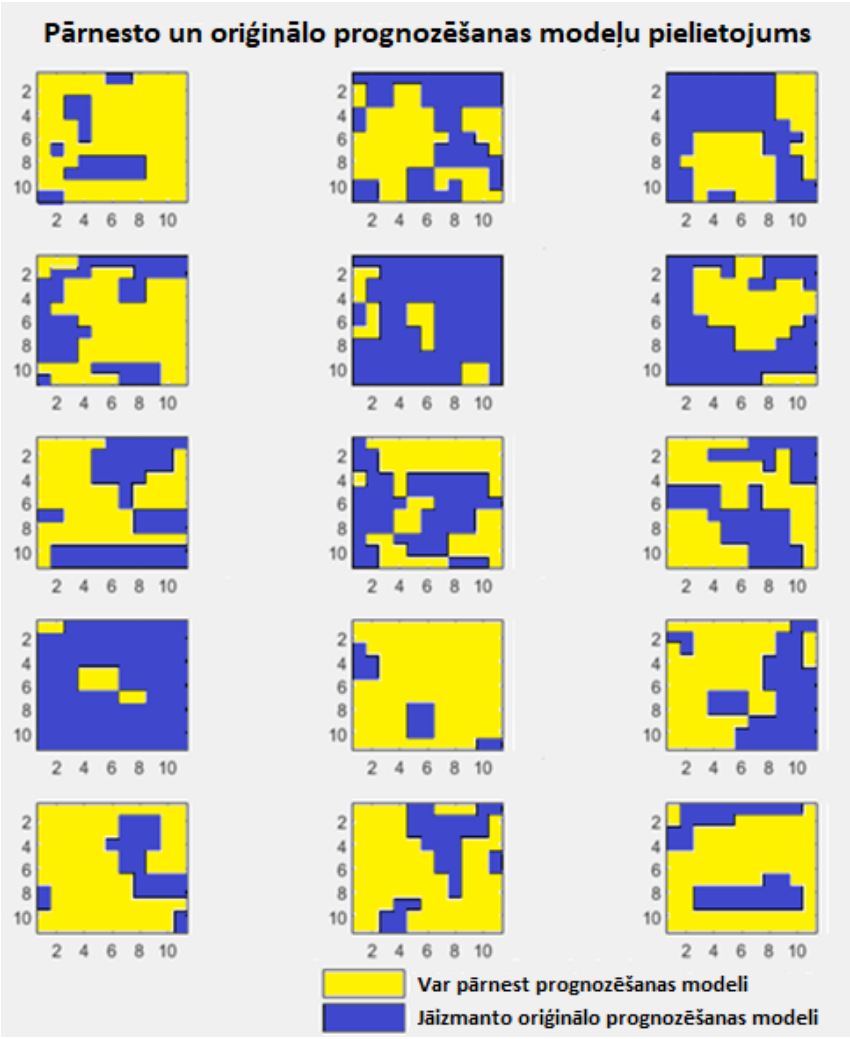
principiālo komponentu analīzi. Visi pirmapstrādei izmantojamie parametri tiek ņemti no iepriekš atrastajiem piemērotajiem datu pirmapstrādes parametriem, kas tika iegūti, veicot pirmapstrādi centrālajā *NDVI* laika rindā. Tad uz centrālās *NDVI* laika rindas apmācītais *LRNN* prognozēšanas modelis tiek izmantots, lai prognozētu katras *NDVI* laika rindas vērtības. Prognozējot katru no šīm 120 kaimiņu pikseļu laika rindām, tiek aprēķinātas zaudējumu funkcijas *RMSE* vērtības starp novēroto un prognozēto laika rindu testa datu kopā atbilstoši vienādojumam (1.3.).

Tālāk uz katras no šīm 120 laika rindām no jauna tiek veikta datu pirmapstrāde: fāžu telpas rekonstrukcija ar laika aiztures metodi, atribūtu atlasīšana, izmantojot pakāpenisko regresijas analīzi, un lineāri nekorelējošu atribūtu kopu iegūšana, izmantojot principiālo komponentu analīzi. Tagad visi datu pirmapstrādes parametri tiek atrasti individuāli uz katras *NDVI* laika rindas atsevišķi. Tālāk uz katras laika rindas datiem tiek apmācīts *LRNN* prognozēšanas modelis, kas tiek izmantots, lai prognozētu katras *NDVI* laika rindas vērtības. Prognozējot katru no šīm 120 kaimiņu pikseļu laika rindām, atkal tiek aprēķinātas zaudējumu funkcijas *RMSE* vērtības starp novēroto un prognozēto laika rindu testa datu kopā atbilstoši 1.3 vienādojumam.

Šādi veikts eksperiments tiek atkārtots 100 reižu, katrā eksperimentā prognozējot 121 *NDVI* laika rindu: centrālo un 120 kaimiņu laika rindas. Piecpadsmit šādos eksperimenta veikšanas gadījumos (no 100 gadījumiem) iegūtie rezultāti parādīti 5.6. attēlā, kas ir binārais attēls, kur ar dzeltenu krāsu (vērtība viens) ir attēloti tie pikseļi, kur var izmantot uz centrālās *NDVI* laika rindas iegūtos datu pirmapstrādes parametrus un prognozēšanas modeli uz konkrētā pikseļa laika rindas.

Savukārt ar zilo krāsu (vērtība nulle) tiek attēloti tie pikseļi, kur ir jāizmanto uz šīs laika rindas individuāli iegūtos datu pirmapstrādes parametrus un prognozēšanas modeli (oriģinālo prognozēšanas modeli). Kā ir redzams 5.6. attēlā, tad katrā no piecpadsmit eksperimenta veikšanas gadījumiem ir daļa laika rindu, kur var izmantot pārnestos datu pirmapstrādes parametrus un prognozēšanas modeli, lai iegūtu prognozes. Arī katrā no 100 eksperimenta gadījumiem ir *NDVI* laika rindas, kuras var veikt pirmapstrādi un prognozēt ar pārnestiem datu pirmapstrādes parametriem un prognozēšanas modeļiem. Tas pierāda darbā izteikto otro hipotēzi.

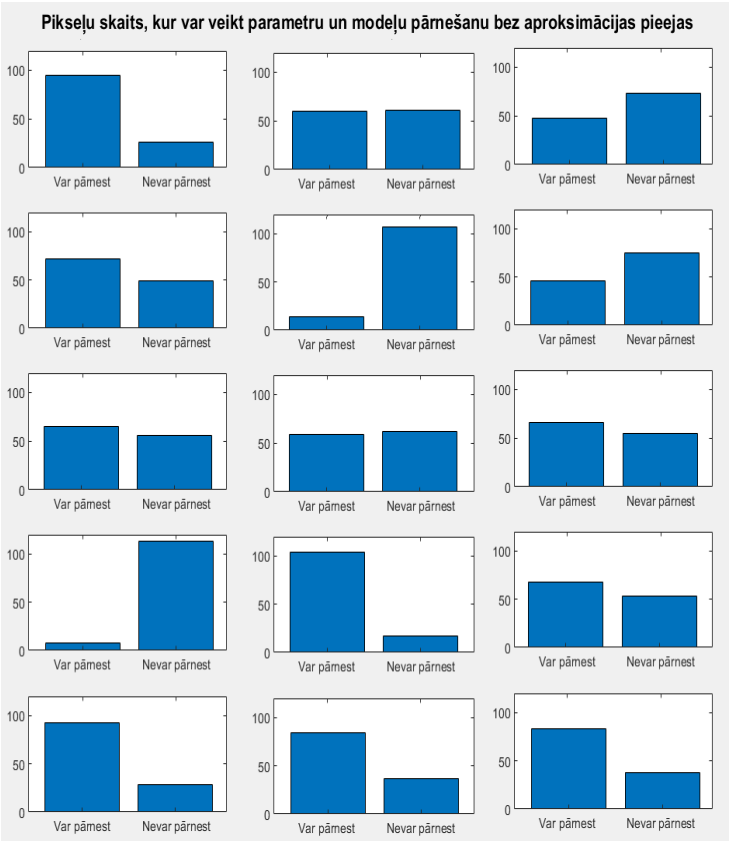
Tādā veidā galvenā priekšrocība šādai datu pirmapstrādes parametru un prognozēšanas modeļu pārnešanai, ir iespēja ātri iegūt *NDVI* vērtību prognozes lieliem zemes virsmas apgabaliem, kur individuāla datu pirmapstrādes parametru meklēšana un piemērota prognozēšanas modeļa iegūšana prasa laiku, kas pārsniedz vienu nedēļu.



5.6. att. Prognozēšanas modeļu lietojums bez aproksimācijas pieejas izmantošanas.

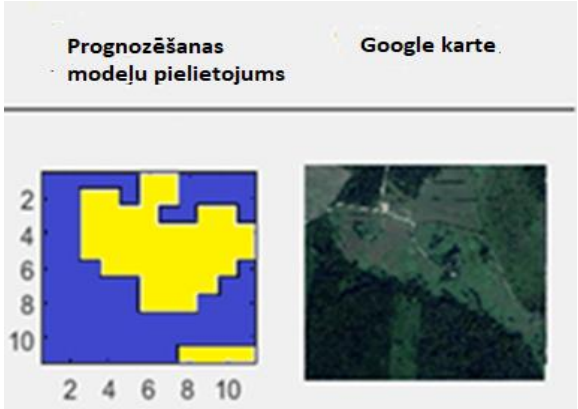
Lai varētu skaitliski novērtēt, uz cik pikseļiem var pārnest iegūtos datu pirmapstrādes parametrus un prognozēšanas modeļus, katram no 5.6. attēlā apskatītajiem piecpadsmit eksperimenta gadījumiem tiek izveidota stabiņu diagramma (skat. 5.7. attēlu).

Diagrammas stabiņš 5.6. attēla kreisajā pusē parāda pikseļu skaitu (no 121 pikseļa), kur var pielietot uz centrālā pikseļa *NDVI* laika rindas iegūtos datu pirmapstrādes parametrus un prognozēšanas modeli. Savukārt stabiņš labajā pusē parāda pikseļu skaitu, kur ir jāizmanto uz šī pikseļa individuāli iegūtos datu pirmapstrādes parametrus un prognozēšanas modeli.



5.7. att. Pikseļu skaits, kur var pārnest parametrus un modeli bez apakšsignāla aproksimācijas pieejas izmantošanas.

Lai varētu noteikt, kuros gadījumos var pārnest datu pirmapstrādes parametrus un prognozēšanas modeļus, ir nepieciešams izstrādāt nosacījumus. Viens no pieņēmumiem ir, ka pārnestie datu pirmapstrādes parametri un prognozēšanas modeļi spēj sniegt labu prognozēšanas precizitāti uz tām laika rindām, kas iegūtas no pikseļiem, kuriem sakrīt zemes virsmas kategorija (mežs, pļava, ūdens u. c.). Šāds pieņēmums ir labi redzams sestajā eksperimenta gadījumā, kas ir parādīts 5.8. attēlā.

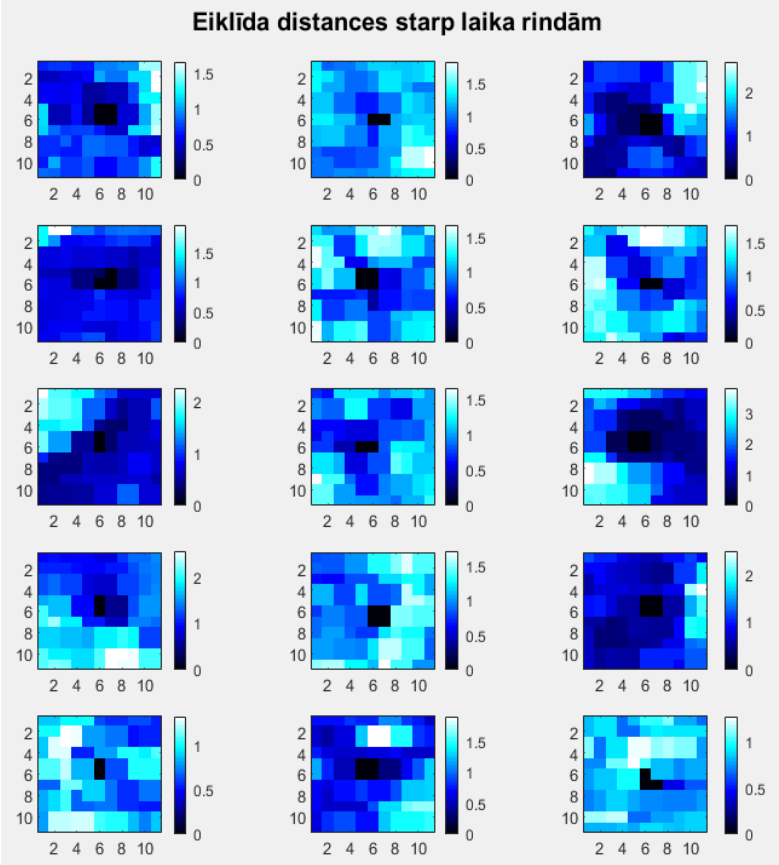


5.8. att. Prognozēšanas modeļu pielietojums sestajā eksperimenta gadījumā un zemes virsmai atbilstoša *Google* karte.

Daudzos gadījumos 5.8. attēlā var veikt pirmapstrādes parametru un prognozēšanas modeļu pārnešanu to pikseļu laika rindām, kurām sakrīt zemes virsmas kategorija ar centrālo pikseli. Tātad, ja centrālajam pikselim atbilst pļavas pārklājums, tad arī uz *NDVI* laika rindām, kas iegūtas no citiem pikseļiem ar pļavas pārklājumu, var tikt izmantoti šie paši datu pirmapstrādes parametri un prognozēšanas modelis.

Lai varētu izmantot zemes virsmas kategorijas kā nosacījumu datu pirmapstrādes parametru un prognozēšanas modeļu pārnešanai, ir nepieciešams iegūt digitālu karti (rastra attēlu), kas saturētu zemes virsmas kategorijas. Darba autors ir veicis pētījumu [89], kur tiek klasificētas zemes virsmas kategorijas. Tomēr šajā pētījumā iegūtā klasifikācijas kopējā precizitāte ir 76,66 %, kamēr nepieciešamā precizitāte būtu daudz augstāka, kas nozīmē, ka ir jāturpina pētījumi šajā virzienā.

Tomēr pat, ja pikseļa zemes virsmas kategorija sakrīt ar centrālā pikseļa zemes virsmas kategoriju, no šī pikseļa iegūtā laika rinda var atšķirties pēc savām īpašībām. Kā laika rindu līdzības mēru var izmantot Eiklīda distanci, un 5.9. attēlā tiek parādītas Eiklīda distances starp centrālā pikseļa laika rindu un pārējo pikseļu laika rindām: jo zemāka vērtība, jo līdzīgākas ir laika rindas.



5.9. att. Eiklīda distances starp centrālā pikseļa laika rindu un kaimiņu laika rindām.

No 5.9. attēla var redzēt, ka no tuvākiem diviem kaimiņu pikseļiem iegūtās laika rindas ir līdzīgas, kur Eiklīda distance ir mazāka par 0,1. Lielākiem pikseļu attālumiem Eiklīda distances vērtība mainās lielāka intervālā, un atkarībā no citiem faktoriem, piemēram, zemes virsmas pārklājuma, var pieņemt mazāku vai lielāku vērtību.

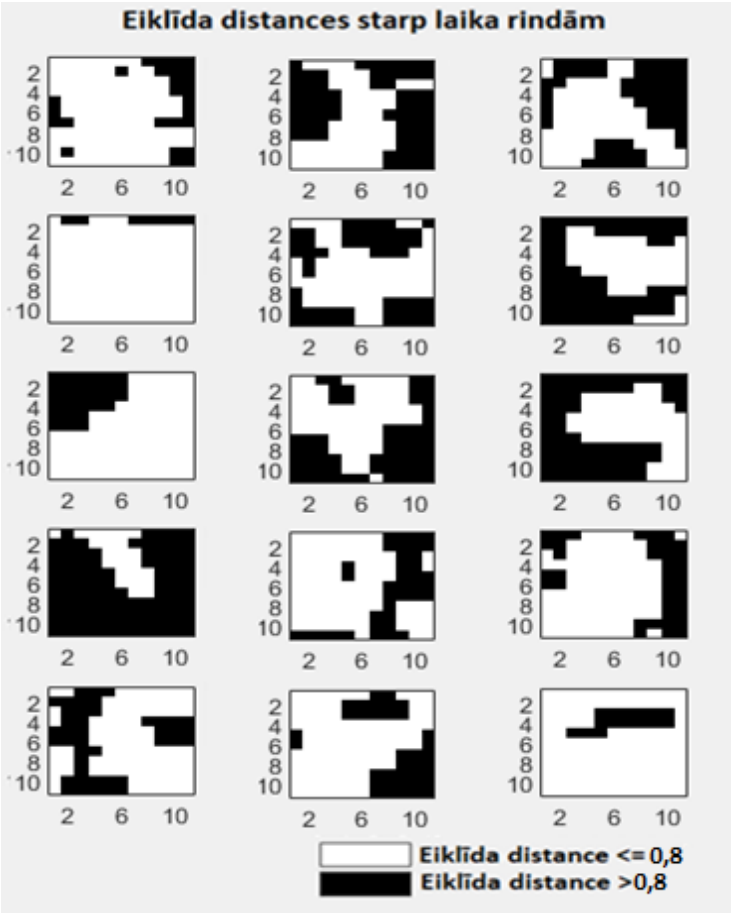
Savukārt 5.10. attēlā tiek parādītas Eiklīda distances, kas ir iedalītas divās kategorijas, atkarībā no izmantotā sliekšņa. Respektīvi, ja Eiklīda distance starp centrālā pikseļa *NDVI* laika rindu un cita pikseļa laika rindu ir mazāka vai vienāda ar 0,8, tad šī pikseļa vērtība ir viens, un tas tiek attēlots ar baltu krāsu, pretējā gadījumā pikseļa vērtība ir nulle, un tas tiek attēlots ar melnu krāsu.

Ja salīdzina pikseļu vērtības 5.10. attēlā un 5.6. attēlā, tad var tikt atrasta likumsakarība, ka to pikseļu laika rindu prognozēšanai, kur Eiklīda distance starp šo laika rindu un centrālo laika rindu ir mazāka vai vienāda ar 0,8, daudzos gadījumos var tikt izmantoti no centrālās laika rindas iegūtie datu pirmapstrādes parametri un prognozēšanas modelis. Tātad nosacījums, lai varētu pārnest datu pirmapstrādes parametrus un prognozēšanas modeli, var tikt definēts šādi: Eiklīda distancei starp centrālo laika rindu un tā pikseļa laika rindu, kuru ir nepieciešams prognozēt, ir jābūt mazākai vai vienādai ar 0,8.

Tomēr, izmantojot Eiklīda distanci ar slieksni, daudzos gadījumos, kur varētu izmantot datu pirmapstrādes parametru un prognozēšanas modeļu pārnešanu, tiek iegūta Eiklīda distances vērtība, kas pārsniedz slieksni, šajā gadījumā vērtību 0,8. Un pretēji, tiem pikseļiem, kuru laika rindām būtu jāizmanto oriģinālais prognozēšanas modelis kopā ar datu pirmapstrādes parametriem, Eiklīda distance ir mazāka par noteikto slieksni.

Tādēļ ir jāturpina pētījumi piemērotu nosacījumu definējuma atrašanā, apskatot gan citus laika rindu līdzības mērus (korelācijas koeficients, Mahalanobisa distance, Minovska distance, u. c.), gan kombinējot vairākus faktorus kopā, piemēram, apvienojot dažādus laika rindu līdzības mērus ar zemes virsmas kategorijām. Kā arī varētu būt nepieciešama *NDVI* laika rindu klasterizācija. Būtu jāpārbauda, vai viena klasterī esošām *NDVI* laika rindām var pielietot vienus un tos pašus datu pirmapstrādes parametrus un prognozēšanas modeli.

Ir jāņem vērā arī fakts, ka visi iegūtie datu pirmapstrādes parametri un prognozēšanas modeļi ir iegūti no izstrādātās prognozēšanas sistēmas *NDVI FS* bez *VMD* apakšsignāla aproksimācijas pieejas izmantošanas.



5.10. att. Eiklīda distances ar slieksni 0,8.

Bet kā parādīja ceturtais un piektais eksperiments, kuri ir aprakstīti 5.5.1. un 5.5.2. punktos, tad labāku prognozēšanas precizitāti uzrāda *NDVI FS* sistēma ar *VMD* apakšsignāla aproksimācijas pieejas izmantošanu. Tādēļ ir nepieciešams novērtēt datu pirmapstrādes parametru un prognozēšanas modeļu pārnešanu, veicot prognozēšanu ar *NDVI FS* sistēmu, kas izmanto *VMD* apakšsignāla aproksimācijas pieeju, kas tiek izdarīts 5.5.4. punktā.

5.5.4. Prognozēšana ar modeļu pārnešanu un ar aproksimācijas pieeju

Septītā eksperimenta ietvaros *NDVI* laika rindu prognozēšana tiek veikta, izmantojot izstrādāto sistēmu *NDVI FS* ar *VMD* apakšsignāla aproksimācijas pieejas izmantošanu, kā tas ir aprakstīts 5.5.2. punktā.

Uz *NDVI* laika rindas tiek veikta datu pirmapstrāde, tiek saglabāti visi nepieciešamie datu pirmapstrādes parametri, un tiek iegūts piemērots prognozēšanas modelis, kas uz apmācībā izmantotās *NDVI* laika rindas validācijas datu kopas sniedz mazāko *RMSE* zaudējumu funkcijas vērtību.

Satelītattēlā tiek novilkts 11x11 pikseļu režģis ar rādiusu $r = 5$, kur bez centrālā pikseļa, uz kura tika veikta apmācība, vēl ir 120 pikseļi. Katram no šiem kaimiņu pikseļiem tiek iegūtas *NDVI* laika rindas, uz kuram tiek veikta datu pirmapstrāde un prognozēšana:

- vispirms katrai no šīm 120 laika rindām, izmantojot modificēto variāciju apakšsignālu dekompozīcijas metodi, tiek iegūts apakšsignāls;
- izstrādātā apakšsignāla aproksimācijas pieeja tiek izmantota, lai tuvināti aprēķinātu apakšsignālu vērtības katrā laika solī;
- gan oriģinālajai *NDVI* laika rindai, gan aproksimētajām apakšsignālam tiek veikta fāžu telpas rekonstrukcija ar laika aiztures metodi;
- pakāpeniskā regresijas analīze tiek izmantota, lai atlasītu informatīvākos atribūtus, un tiek iegūtas lineāri nekorelējošu atribūtu kopas, izmantojot principiālo komponentu analīzi;
- visi pirmapstrādei izmantojamie parametri tiek paņemti no iepriekš atrastajiem datu pirmapstrādes parametriem, kas tika iegūti, veicot pirmapstrādi uz centrālās *NDVI* laika rindas.
- prognozēšana tiek veikta, izmantojot uz centrālās *NDVI* laika rindas apmācīto *LRNN* prognozēšanas modeli.

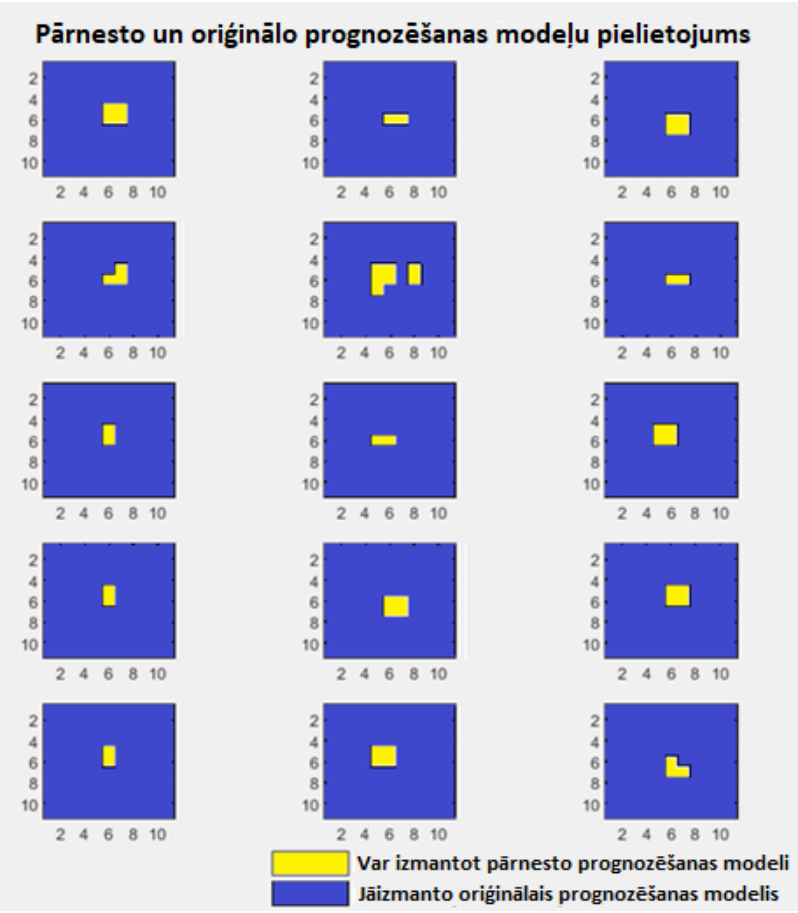
Prognozējot katru no šīm 120 kaimiņu pikseļu laika rindām, tiek aprēķinātas zaudējumu funkcijas *RMSE* vērtības starp novēroto un prognozēto laika rindu testa datu kopā atbilstoši vienādojumam (1.3.). Tālāk uz katras no šīm 120 laika rindām no jauna tiek veikta datu pirmapstrāde:

- apakšsignāla iegūšana no modificētās variāciju apakšsignālu dekompozīcijas metodes;
- apakšsignāla vērtību tuvinātā aprēķināšana katrā laika solī;
- *NDVI* laika rindas un aproksimētā apakšsignāla fāžu telpu rekonstrukcija ar laika aiztures metodi;
- atribūtu atlasīšana, izmantojot pakāpenisko regresijas analīzi;
- lineāri nekorelējošu atribūtu kopu iegūšana, izmantojot principiālo komponentu analīzi.

Tagad visi datu pirmapstrādes parametri tiek atrasti individuāli uz katras *NDVI* laika rindas atsevišķi. Tālāk uz katras no šīm laika rindām tiek apmācīts *LRNN*

prognozēšanas modelis, kas tiek izmantots, lai prognozētu katras *NDVI* laika rindas vērtības. Prognozējot katru no šīm 120 kaimiņu pikseļu laika rindām, atkal tiek aprēķinātas zaudējumu funkcijas *RMSE* vērtības starp novēroto un prognozēto laika rindu testa datu kopā atbilstoši 1.3 vienādojumam.

Šādi veikts eksperiments tiek atkārtots 100 reižu, katrā eksperimentā prognozējot 121 *NDVI* laika rindu: centrālo un 120 kaimiņu laika rindas. Turklāt visos 100 septītā eksperimenta gadījumos izmantotās laika rindas ir vienas un tās pašas, kas tiek izmantotas sestajā eksperimentā. Piecpadsmit šādos eksperimenta veikšanas gadījumos iegūtie rezultāti parādīti 5.11. attēlā.



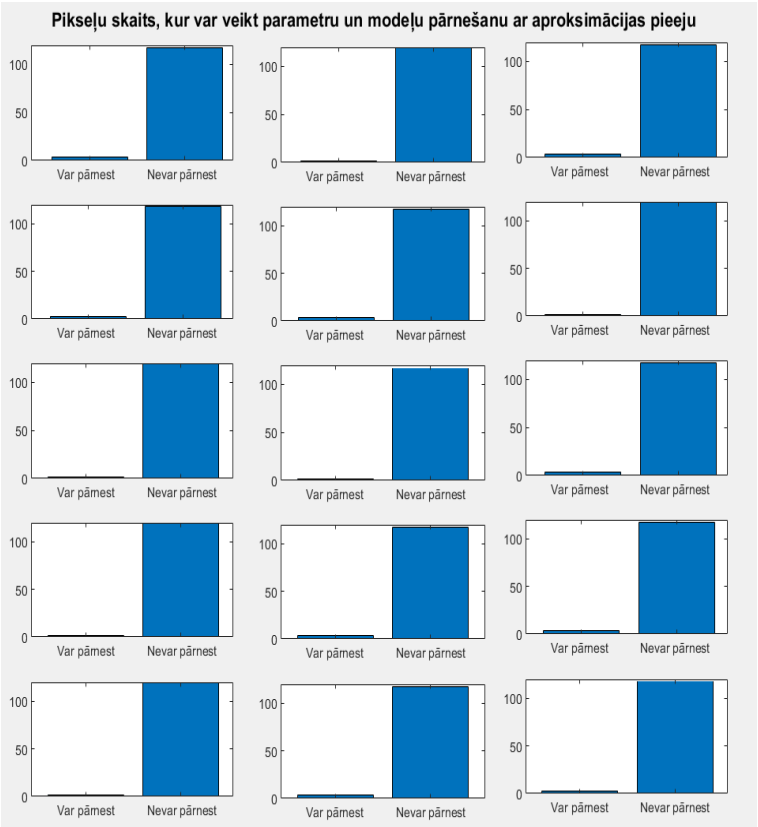
5.11. att. Prognozēšanas modeļu lietojums ar aproksimācijas pieejas izmantošanu.

Ir redzams, ka 5.11. attēls ir binārais attēls, kur ar dzeltenu krāsu (vērtība viens) ir attēloti tie pikseļi, kur var pārnest uz centrālās *NDVI* laika rindas iegūtos datu pirmapstrādes parametrus un prognozēšanas modeli, lai veiktu konkrētā pikseļa laika rindas pirmapstrādi un prognozēšanu. Šo pikseļu laika rindas tiek prognozētas ar līdzvērtīgu precizitāti gan izmantojot individuāli iegūto (oriģinālo), gan pārnesto prognozēšanas modeli. Savukārt ar zilo krāsu (vērtība nulle) tiek attēloti tie pikseļi, kur

ir jāizmanto individuāli iegūtie datu pirmapstrādes parametri un prognozēšanas modelis. Kā ir redzams 5.11. attēlā, tikai tuvākie kaimiņu pikseļi ap centrālo pikseli pieņem vērtību viens.

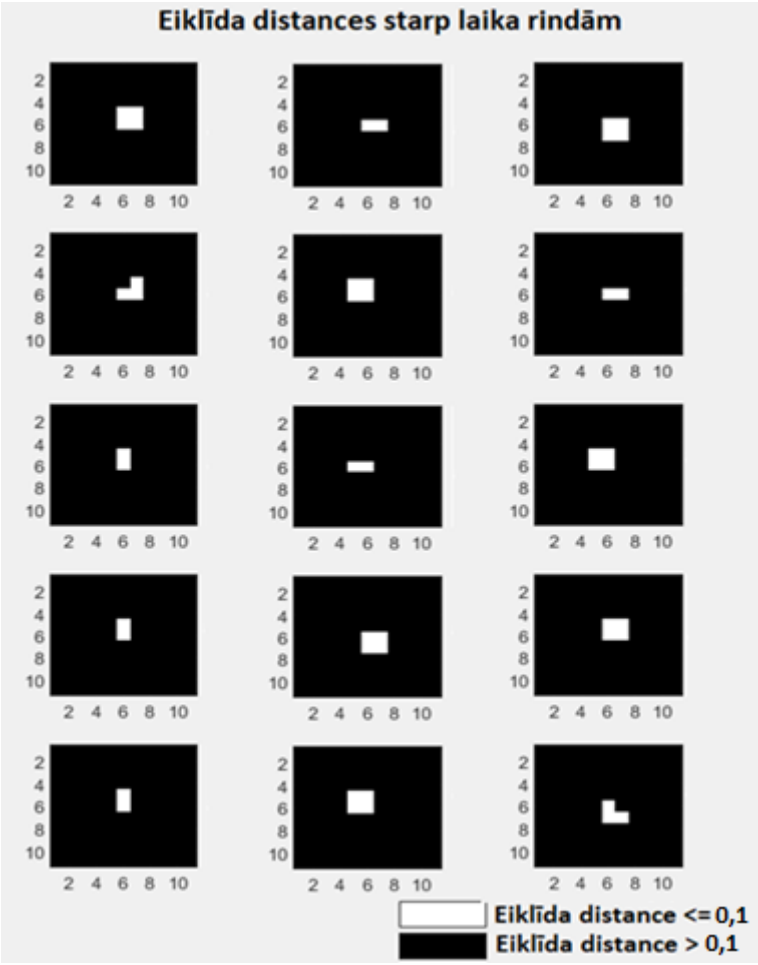
Koeficientu matrica (skat. matricu (3.16.)) tiek iegūta uz centrāla pikseļa *NDVI* laika rindas un tiek izmantota arī, lai tuvināti aprēķinātu no *VMD* metodes iegūto apakšsignālu uz pārējām 120 laika rindām. Savukārt brīvo locekļu vektoru aprēķina pēc vienādojumiem (3.17.) un (3.18.), un tā aprēķināšanai tiek izmantotas centrāla pikseļa *NDVI* laika rindas vērtības. Ja centrāla pikseļa *NDVI* laika rindu aizvieto ar jebkuru no 120 kaimiņu laika rindām, un tās izmanto brīvo locekļu vektora aprēķināšanai, tad tas noved pie lielām izmaiņām atrisinājumā. Un šāda atrisinājuma vērtību padošana ieejā *LRNN* prognozēšanas modelim neveicina prognozēšanas precizitātes pieaugumu.

Stabiņu diagramma (skat. 5.12. attēlu) tiek izveidota, lai parādītu pikseļu skaitu, kur var veikt pirmapstrādes parametru un modeļu pārnešanu. Diagrammas stabiņš 5.12. attēla kreisajā pusē parāda pikseļu skaitu no 121 pikseļa, kur var pielietot uz centrālā pikseļa *NDVI* laika rindas iegūtos datu pirmapstrādes parametrus un prognozēšanas modeli, bet labajā pusē parāda pikseļu skaitu, kur nevar veikt pārnešanu.



5.12. att. Pikseļu skaits, kur var pārnest parametrus un modeli ar apakšsignāla aproksimācijas pieejas izmantošanu.

Līdzvērtīga prognozēšanas precizitāte, pārnesot datu pirmapstrādes parametrus un prognozēšanas modeli, tiek sasniegta tikai tiem pikseļiem, kur Eiklīda distance starp šo pikseļu laika rindu un centrālo laika rindu ir zemāka par noteiktu slieksni, kā parādīts 5.13. attēlā. Kā var redzēt 5.13. attēlā, ir izvēlēts slieksnis 0,1 un, ja salīdzina ar 5.11. attēlu, var redzēt, ka pilnībā sakrīt pikseļu vērtības.



5.13. att. Eiklīda distances ar slieksni 0,1.

Tātad nosacījums var tikt definēts šādi: ja Eiklīda distance starp centrālā pikseļa laika rindu un cita pikseļa laika rindu ir mazāka vai vienāda ar 0,1, tad no *NDVI FS* sistēmas, kas izmanto *VMD* apakšsignāla aproksimācijas pieeju, iegūtos datu pirmapstrādes parametrus un prognozēšanas modeļus var izmantot šīs laika rindas pirmapstrādei un prognozēšanai.

Šie rezultāti arī apstiprina darbā izvirzīto otro hipotēzi. Ja tiek izmantota aproksimācijas pieeja, tad nosacījums labi apraksta gadījumus, kad var izmantot prognozēšanas modeļu un datu pirmapstrādes parametru pārnešanu.

5.6. Prognozēšanas precizitātes rezultātu apkopojums

Piektajā nodaļā tiek veikti pieci eksperimenti, lai savstarpēji salīdzinātu vienkāršā slīdošā vidējā (*SMA*), diskrētā laika, nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēžu, automātiski regresējoša integrētā slīdošā vidējā (*ARIMA*), un izstrādātās prognozēšanas sistēmas (izmantojot no modificētās *VMD* metodes iegūtā apakšsignāla aproksimācijas pieeju, un bez šīs pieejas izmantošanas) sasniegto prognozēšanas precizitāti. Ar visām šīm metodēm iegūtās zaudējumu funkciju: saknes vidējās kvadrātiskās kļūdas (*RMSE*), virzienu simetrijas (*DS*) un pielāgotā determinācijas koeficienta R^2_{adj} vidējās vērtības testa datu kopās tiek dotas 5.8. tabulā.

5.8. tabula

Zaudējumu funkciju vidējās vērtības

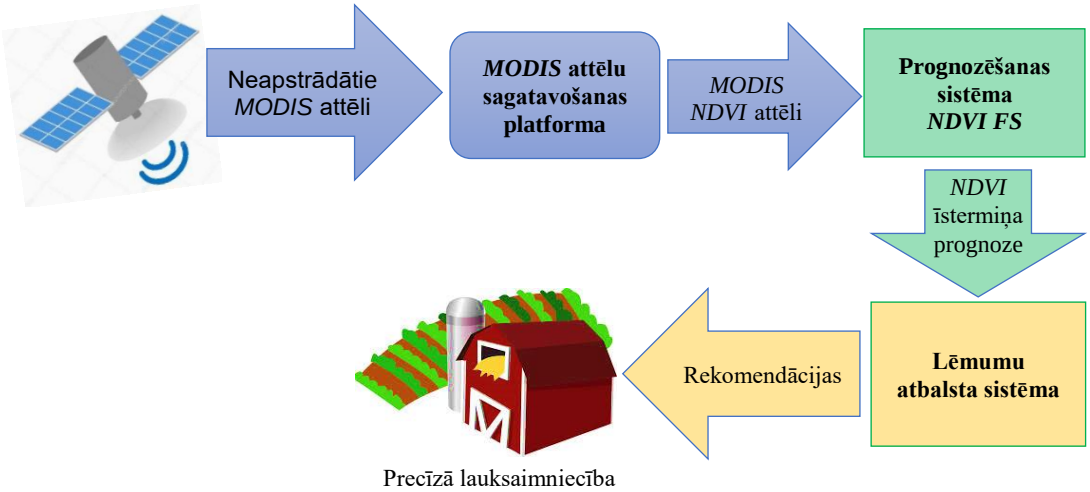
Prognozēšanas metode vai sistēma	***RMSE***	***DS***	$\boldsymbol{R^2_{adj}}$
Vienkāršais slīdošais vidējais	0,0442	93,40 %	0,94
Diskrētā laika, nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēde	0,0214	83,76 %	0,90
Autoregresīvais integrētais slīdošais vidējais	0,0108	93,85 %	0,96
Izstrādāta prognozēšanas sistēma *NDVI FS* bez *VMD* apakšsignāla aproksimācijas pieejas	0,0012	97,61 %	**0,99**
Izstrādāta prognozēšanas sistēma *NDVI FS* ar *VMD* apakšsignāla aproksimācijas pieeju	**0,0009**	**98,80 %**	**0,99**

Prognozēšanas metodes un sistēmas 5.8. tabulā ir sakārtotas proporcionāli vidējo *RMSE* vērtību kritumam, respektīvi, prognozēšanas precizitātes pieaugumam. Viszemāko vidējo *RMSE* vērtību sasniedz vienkāršā slīdošā vidējā metode, iegūstot vidējo vērtību 0,0442. Otro sliktāko *RMSE* vērtību 0,0214 uzrāda diskrētā laika, nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēdes. Tomēr, salīdzinot divu citu zaudējumu funkciju: *DS* un pielāgotā R^2_{adj} vidējās vērtības, tad vienkāršā slīdošā vidējā metode ($DS = 93,40\%$ un pielāgotais $R^2 = 0,94$) uzrāda labākus rezultātus, salīdzinot ar Markova ķēdēm ($DS = 83,76\%$ un pielāgotais $R^2_{adj} = 0,90$). Trešajā vietā ($RMSE = 0,0108, DS = 93,85\%$ un pielāgotais $R^2_{adj} = 0,96$) ir automātiski regresējoša integrētā slīdošā vidējā metode. Divus labākos rezultātus pēc visām trim

zaudējumu funkciju: *RMSE*, DS, un pielāgotā R^2_{adj} vidējām vērtībām uzrāda autora izstrādātā prognozēšanas sistēma *NDVI FS*. Respektīvi, otrajā vietā pēc prognozēšanas precizitātes atrodas *NDVI FS* bez *VMD* apakšsignāla aproksimācijas pieejas izmantošanas ($RMSE = 0,0012$, $DS = 97,61\%$ un pielāgotais $R^2_{adj} = 0,99$). Bet augstāko prognozēšanas precizitāti uzrāda *NDVI FS* ar *VMD* apakšsignāla aproksimācijas pieejas izmantošanu ($RMSE = 0,0009$, $DS = 98,80\%$ un pielāgotais $R^2_{adj} = 0,99$). Šie rezultāti (skat. 5.8. tabulu) parāda autora izstrādātās sistēmas augstu precizitāti *NDVI* laika rindu īstermiņa prognozēšanā. Šajos piecos eksperimentos iegūto *RMSE* vērtību sadalījumi (skat. **4**. pielikumu) uzskatāmi parāda, cik bieži katra no iegūtajām vērtībām ir sastopamas 100 katra eksperimenta gadījumos.

5.7. Prognozēšanas sistēmas *NDVI FS* pielietošana

Izstrādātā prognozēšanas sistēma *NDVI FS* ir tas posms precīzajā lauksaimniecībā, kas nodrošina ieejas datus (*NDVI* īstermiņa prognozi) lēmumu atbalsta sistēmai (skat. 5.14. att.). Sistēma *NDVI FS* saņem *MODIS NDVI* attēlus ieejā no kādas *MODIS* attēlu sagatavošanas platformas, kas šajā darbā ir Vīnes Dabas resursu un dabaszinātņu universitātes *MODIS* veģetācijas indeksu laika rindu datu apstrādes platforma [106]. Izejā sistēma *NDVI FS* izsniedz lietotāja (piemēram, lauksaimnieka) izvēlētās *NDVI* laika rindas īstermiņa prognozi, kas nonāk ieejā lēmumu atbalsta sistēmai. Lēmumu atbalsta sistēma izmanto *NDVI* indeksa prognozi, lai pieņemtu noteiktus lēmumus un izsniegtu rekomendācijas lauksaimniekam.



5.14. att. Izstrādātā prognozēšanas sistēma *NDVI FS* precīzajā lauksaimniecībā.

Kā tika apskatīts 1.7. apakšnodaļā, precīzajā lauksaimniecībā ir divi galvenie *NDVI* indeksa prognozes pielietojumu veidi:

- *NDVI* prognozes salīdzināšana ar iepriekšējo gadu *NDVI* vērtībām šajā pašā datumā. Atkarībā no starpības lieluma un virziena (pozitīvs vai negatīvs) starp prognozēto *NDVI* vērtību un katru no iepriekšējo gadu *NDVI* vērtībām, var tikt pieņemts lēmums par minerālmēslojuma un citu barības vielu apjoma palielināšanu noteiktam zemes nogabalam, ļaujot iegūt labāku ražu. Vai arī tieši pretēji – samazināšanu, kas ļauj samazināt izdevumus, kā arī tiek nodarīts mazāks kaitējums apkārtējai videi.
- *NDVI* prognoze tiek izmantota dažādos kultūraugu ražības modeļos. Tas ļauj iegūt gan prognozēto ražas apjomu nākamajam laika periodam (piemēram, *MODIS NDVI* attēliem tās ir 7 dienas uz priekšu), gan arī aprēķināt sagaidāmos ienākumus. Savukārt tas ļauj pieņemt lēmumu par piemēroto ražas novākšanas datumu, aprēķinot, kad būs sagaidāmi maksimālie ienākumi.

Ražas prognozēšana, kā arī sagaidāmo ienākumu prognozes ļauj lauksaimniekam savlaicīgi noslēgt produkcijas piegādes līgumus un līdz ar to veiksmīgi konkurēt lauksaimniecības tirgus nozarē. Piemēram, pētījumā [98] izmanto eksponenciālo vienādojumu (5.1.), lai aprakstītu graudu ražas apjomu y, ko izsaka tonnās uz hektāru:

$$y = 0.936e^{3.3525x}, \qquad (5.1.)$$

kur x – normalizētā veģetācijas indeksa vērtība.

Ja vienādojumā (5.1.) ievieto prognozēto *NDVI* vērtību, tad tiek iegūta ražas prognoze vienu nedēļu uz priekšu. Zemāk redzamajā 5.9. tabulā ir dotas *NDVI* vērtības 2015. gada jūlija mēneša pēdējai nedēļai, augusta mēnesim un septembra pirmajai nedēļai. No 5.9. tabulas ir redzams, ka izstrādātā sistēma *NDVI FS* ir precīzi ļāvusi noteikt ražas novākšanas datumu, kamēr vienkāršais slīdošais vidējais ir parādījis, ka *NDVI* vērtību kritums ir sagaidāms tikai nedēļu vēlāk.

Lauksaimnieks var aprēķināt sagaidāmo ražu un ienākumus uz katru nākamo nedēļu, līdz tiek prognozēts *NDVI* vērtību kritums. Jūlija mēnesī (laika solī $t-5$), kad tuvojas labības novākšanas laiks, lauksaimnieks iegūst ražas prognozi pašreizējai nedēļai, izmantojot patieso *NDVI* vērtību un nākamajai nedēļai (ar laika soli $t-4$).

5.9. tabula

Novērotās un prognozētās *NDVI* vērtības

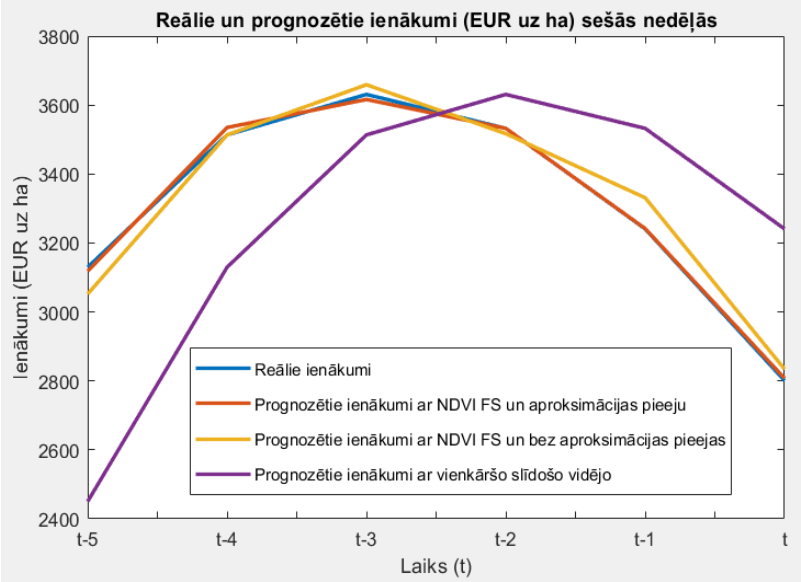
Nr.	Datu avots	Jūlija nedēļa (*t-5*)	Augusta nedēļa (*t-4*)	Augusta nedēļa (*t-3*)	Augusta nedēļa (*t-2*)	Augusta nedēļa (*t-1*)	Septembra nedēļa (*t*)
1.	Novērotā *NDVI* vērtība	0,8257	0,8602	0,8700	0,8618	0,8361	0,7926
2.	Izstrādātā prognozēšanas sistēma *NDVI FS* ar aproksimācijas pieeju	0,8246	0,8620	0,8688	0,8618	0,8362	0,7933
3.	Izstrādātā prognozēšanas sistēma *NDVI FS* bez aproksimācijas pieejas	0,8182	0,8602	0,8723	0,8605	0,8443	0,7961
4.	Vienkāršais slīdošais vidējais	0,7527	0,8257	0,8602	0,8700	0,8618	0,8361

Lauksaimnieks iegūst ražas prognozi, izmantojot izstrādāto sistēmu *NDVI FS* ar apakšsignāla aproksimācijas pieejas izmantošanu, bez apakšsignāla aproksimācijas pieejas izmantošanas un lietojot vienkāršo slīdošo vidējo. Ja tiek prognozēts ražas apjoma kāpums, tad vēl netiek veikti ražas novākšanas darbi. Bet lauksaimnieks, izmantojot nākamās nedēļas ražas prognozi, var slēgt piegādes līgumus ar labības pircējiem par noteikta ražas apjoma piegādi, tādējādi apsteidzot konkurentus. Pēc nedēļas tiek prognozēta raža laika solim $t-3$, un joprojām tiek prognozēts ražas apjoma kāpums. Tātad raža vēl nav jānovāc. Lauksaimnieks var turpināt slēgt līgumus ar pircējiem, iekļaujot arī jaunajā prognozē iegūto ražas apjomu.

Nākamajā nedēļā (laika solī $t-3$), izmantojot izstrādāto sistēmu *NDVI FS* (gan ar, gan bez apakšsignāla aproksimācijas pieejas), tiek iegūta prognoze ar ražas apjoma kritumu $t-2$ nedēļai. Šajā brīdī ir jānovāc raža. Tomēr, skatoties vienkāršā slīdošā vidējā izsniegto prognozi 5.9. tabulā, joprojām tiek prognozēts ražas apjoma kāpums. Pēc šīs prognozes lauksaimniekam vēl nebūtu jānovāc raža, bet tikai nākamajā nedēļā, un tad noved pie ienākumu samazināšanās.

Reālie un prognozētie ienākumi tiek aprēķināti, izmantojot *NDVI* reālās vērtības un atbilstošās prognozes, kas iegūtas no izstrādātās sistēmas *NDVI FS* ar un bez apakšsignāla aproksimācijas pieejas izmantošanas, kā arī izmantojot vienkāršo slīdošo

vidējo (skat. 5.9. tabulu). Reālās un prognozētās *NDVI* vērtības tiek izmantotas, lai iegūtu graudu ražas apjomu (t/ha) pēc vienādojuma (5.1.). Graudu ražas apjoms tiek sareizināts ar cenu: 204,91 *EUR* par tonnu, iegūstot sagaidāmo ienākumu prognozes (skat. 5.15. att.).



5.15. att. Reālie un prognozētie ienākumi no graudu ražas apjoma.

Prognozēto ienākumu kļūda uz hektāru $t-3$ nedēļā izmantojot izstrādāto prognozēšanas sistēmu *NDVI FS* ar un bez aproksimācijas pieejas, kā arī vienkāršo slīdošo vidējo, ir šāda:

- izstrādātā sistēma *NDVI FS* ar aproksimācijas pieeju: prognozēto ienākumu kļūda uz hektāru – par 14,23 *EUR* mazāk nekā reālie ienākumi;
- izstrādātā sistēma *NDVI FS* bez aproksimācijas pieejas: prognozēto ienākumu kļūda uz hektāru – par 27,43 *EUR* vairāk nekā reālie ienākumi;
- vienkāršais slīdošais vidējais: prognozēto ienākumu kļūda uz hektāru – par 96,10 *EUR* vairāk nekā reālie ienākumi.

Tātad, izmantojot sistēmu *NDVI FS* ar aproksimācijas pieeju, prognozētie ienākumi uz katru hektāru ir par 14,23 *EUR* mazāki nekā reālie ienākumi, un tas ļauj šo sistēmu praktiski izmantot precīzajā lauksaimniecībā.

5.8. Kopsavilkums un secinājumi

Izveidota *NDVI FS*, ar kuru iespējams noteikt *NDVI* laika rindas jaunas vērtības prognozi nākamajai nedēļai. Iegūto prognozi izmanto precīzās lauksaimniecības

speciālists, kurš var pieņemt lēmumu par nepieciešamajiem pasākumiem kultūraugu audzēšanā, piemēram, palielināt vai samazināt barības vielu daudzumu, kas tiek izmantots kultūraugu barošanai. Eksperimenti parāda, ka izveidotā *NDVI FS* spēj apstrādāt un ar augstu precizitāti prognozēt *NDVI* laika rindas. Iegūtie rezultāti – *NDVI* laika rindas nākamās nedēļas vērtības skaitliskā prognoze – ir viegli interpretējami sistēmas lietotājam. Galvenie secinājumi par izveidoto normalizēta veģetācijas indeksa laika rindu prognozēšanas pieeju, izstrādāto *NDVI FS* sistēmu un veiktajiem eksperimentiem ir šādi.

1. Prognozēšanas sistēma *NDVI FS* uzrāda labāku precizitāti *NDVI* laika rindu prognozēšanā, salīdzinot ar vienkāršā slīdošā vidējā (*SMA*), nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēžu un automātiski regresējošā integrētā slīdošā vidējā (*ARIMA*) sasniegto precizitāti.
2. Izstrādātā sistēma *NDVI FS* bez aproksimācijas pieejas izmantošanas daudzos gadījumos spēj sasniegt līdzvērtīgu prognozēšanas precizitāti, izmantojot uz citas *NDVI* laika rindas iegūtos datu pirmapstrādes parametrus un prognozēšanas modeli.
3. Izstrādātā sistēma *NDVI FS* bez aproksimācijas pieejas izmantošanas var tikt pielietota, ja nepieciešams prognozēt *NDVI* vērtības lielam zemes virsmas apgabalam, kur ir ierobežots skaitļošanas laiks līdz vienai nedēļai un kur ir pieļaujama neliela prognozēšanas precizitātes samazināšanās.
4. Lai varētu veikt no *NDVI FS* sistēmas iegūto datu pirmapstrādes parametru un prognozēšanas modeļu pārnešanu un izmantošanu citas *NDVI* laika rindas pirmapstrādei un prognozēšanai, Eiklīda distancei starp apmācībā izmantoto laika rindu un otro laika rindu ir jābūt mazākai vai vienādai ar noteiktu slieksni.

Izveidotās *NDVI FS* sistēmas rezultātu interpretācija, ir viegli uztverama un saprotama sistēmas lietotājam, jo skaidri parāda prognozējamās *NDVI* laika rindas vērtību nākamajam laika periodam, kas šajā darbā ir septiņas dienas.

REZULTĀTI UN SECINĀJUMI

Promocijas darbā ir izstrādāta *NDVI* indeksa laika rindu prognozēšanas sistēma *NDVI FS*, kas nodrošina īstermiņa veģetācijas izmaiņu dinamikas prognozes, kas ir vitāli svarīgas precīzajā lauksaimniecībā. Izstrādātā prognozēšanas sistēma tiek eksperimentāli novērtēta, izmantojot autora piedāvāto *VMD* apakšsignāla aproksimācijas pieeju, un bez šīs pieejas, lai pārbaudītu pirmo izvirzīto hipotēzi. Lai pārbaudītu otro izvirzīto hipotēzi, tiek veikti modeļu pārnešanas eksperimenti, uz vienas *NDVI* laika rindas apmācīto prognozēšanas modeli kopā ar datu pirmapstrādes parametriem, izmantojot kaimiņu pikseļu *NDVI* laika rindu prognozēšanai, neveicot jaunu apmācību. Atrisinot promocijas darbā izvirzītos **uzdevumus**, iegūtie galvenie **rezultāti** ir šādi.

1. Veikta zinātniskās literatūras analīze par normalizēta veģetācijas indeksa laika rindu prognozēšanu. Definētas prasības un pamatota prognozēšanas metodes izvēle izstrādājamajai sistēmai.
2. Veikta zinātniskās literatūras analīze par laika rindu prognozēšanu, izmantojot dažādas signālu dekompozīcijas metodes, ko lieto frekvenču analīzei. Pamatota dekompozīcijas metodes izvēle *NDVI* laika rindu prognozēšanas uzdevumā un no tās iegūto apakšsignālu aproksimācijas nepieciešamība.
3. Izstrādāta no modificētās variāciju apakšsignālu dekompozīcijas metodes iegūtā apakšsignāla aproksimācijas pieeja, kas ļauj iegūt tuvinātas apakšsignāla vērtības visiem tiem laika soļiem, kuriem ir pieejami vēsturiskie normalizēta veģetācijas indeksa novērojumi.
4. Izstrādāta normalizēta veģetācijas indeksa laika rindu prognozēšanas sistēma, kura balstās uz specializētu metožu un pieeju kopumu. Tas ļauj paaugstināt prognozēšanas precizitāti, neizmantojot papildu ieejas datus, piemēram, gaisa temperatūras, nokrišņu daudzuma un zemes virsmas kategoriju datus, kā arī citus veģetācijas indeksus. Viss sistēmā implementētais datu pirmapstrādes process, *LRNN* modeļa apmācība un laika rindas prognozēšana ir pilnībā automatizēta.
5. Veikta izstrādātās sistēmas *NDVI FS* novērtēšana, tās precizitāte gan izmantojot, gan neizmantojot apakšsignāla aproksimācijas pieeju ir salīdzināta ar vienkāršā slīdošā vidējā, diskrētā laika, nepārtrauktu stāvokļu kopas *m*-tās

kārtas Markova ķēžu un automātiski regresējoša integrētā slīdošā vidējā sasniegto precizitāti *NDVI* laika rindu prognozēšanā.

6. Izstrādāta pieeja apmācīta prognozēšanas modeļa un atbilstošo datu pirmsapstrādes parametru pārnešanai uz teritoriju, kur nav pieejami prognozēšanas modeļi. Iegūtā prognozēšanas precizitāte daudzos gadījumos ir līdzvērtīga ar rezultātiem, ko sniedz uz konkrētās normalizēta veģetācijas indeksa laika rindas apmācītais prognozēšanas modelis. Modeļu pārnešana palīdz ekonomēt laiku, kas nepieciešams, lai katrai laika rindai atsevišķi iegūtu datu pirmapstrādes parametrus un apmācītu prognozēšanas modeli.

Promocijas darba izstrādāšanas gaitā ir iegūti šādi **secinājumi** par izveidoto normalizēta veģetācijas indeksa laika rindu prognozēšanas sistēmu *NDVI FS*.

1. Normalizēta veģetācijas indeksa laika rindu prognozēšanas pētījumos nepietiekami tiek izmantotas datu pirmapstrādes metodes, kas neļauj sasniegt augstu prognozēšanas precizitāti.
2. Izstrādātā prognozēšanas sistēma veic normalizēta veģetācijas indeksa laika rindas jaunās vērtības īstermiņa prognozi, ieejā izmantojot tikai *NDVI* vēsturiskās vērtības.
3. Izstrādātā pieeja aproksimē no variāciju apakšsignālu dekompozīcijas metodes iegūto apakšsignālu jebkurā laika solī, kur ir pieejamas vēsturiskās normalizēta veģetācijas indeksa laika rindas vērtības, un tas ļauj izmantot šo tuvināti aprēķināto apakšsignālu jaunu *NDVI* laika rindas vērtību prognozēšanai.
4. Aproksimētā apakšsignāla precizitāte ir atkarīga gan no otrās nehomogēnās lineāro vienādojumu sistēmas koeficientu matricas nosacītības skaitļa lieluma, gan no lineāro vienādojumu sistēmu atrisinājumu kļūdu lieluma.
5. Prognozēšanas precizitāte, izmantojot izstrādāto sistēmu un autora piedāvāto apakšsignāla aproksimācijas pieeju, ir augstāka nekā, neizmantojot šo pieeju, kas apstiprina darbā izvirzīto pirmo hipotēzi.
6. Prognozēšanas sistēmas *NDVI FS* sasniegtā normalizēta veģetācijas indeksa laika rindu prognozēšanas precizitāte, gan izmantojot, gan neizmantojot autora piedāvāto aproksimācijas pieeju, ir augstāka nekā precizitāte, kas ir sasniegta, prognozējot *NDVI* laika rindas ar vienkāršā slīdošā vidējā modeli, diskrētā laika, nepārtrauktu stāvokļu kopas *m-tās* kārtas Markova ķēdēm un autoregresīvo slīdošā vidējā modeli.

7. Izmantojot prognozēšanas sistēmu *NDVI FS*, un, veicot apmācību uz kādas *NDVI* laika rindas, var tikt iegūti datu pirmapstrādes parametri un prognozēšanas modelis, ko var izmantot citas *NDVI* laika rindas prognozēšanai ar līdzvērtīgu precizitāti, salīdzinot ar precizitāti, ko var sasniegt, prognozējot šo laika rindu ar individuāli iegūtiem datu pirmapstrādes parametriem un prognozēšanas modeli, ja Eiklīda distance starp šīm abām laika rindām ir mazāka vai vienāda ar kādu noteiktu slieksni. Tas apstiprina darbā izvirzīto otro hipotēzi.
8. Izstrādātā sistēma *NDVI FS* bez aproksimācijas pieejas izmantošanas var tikt lietota, ja nepieciešams prognozēt normalizēta veģetācijas indeksa laika rindu vērtības lielam zemes virsmas apgabalam salīdzinoši īsā laika posmā, un kur ir pieļaujama neliela prognozēšanas precizitātes samazināšanās.
9. Izstrādātā sistēma *NDVI FS* ar aproksimācijas pieejas izmantošanu var tikt lietota, ja nepieciešamas normalizēta veģetācijas indeksa laika rindu prognozes ar augstāku precizitāti, bet salīdzinoši nelieliem zemes virsmas apgabaliem.

Turpmākie pētījumi ir saistīti ar datu pirmapstrādes parametru un prognozēšanas modeļu pārnešanas pieejas attīstīšanu, veicot precīzāku nosacījumu definēšanu, kad drīkst veikt šādu pārnešanu, kā arī uzlabojot aproksimācijas pieeju, paaugstinot tuvināti aprēķināto apakšsignāla vērtību precizitāti.

IZMANTOTĀS LITERATŪRAS SARAKSTS

1. Adhikari, R., Agrawal, R. K. *An Introductory Study on Time series Modeling and Forecasting*. Germany: LAP Lambert Academic Publishing, 2013. 76 pp. ISBN 978-3-659-33508-2. Available from: doi: 10.13140/2.1.2771.8084.

2. ADVANPIX. *Multiprecision Computing Toolbox* [online]. [viewed 2 March 2019]. Available from: https://www.advanpix.com/.

3. Ahmed, D., Elkettan, Y., Kasem, A. Application of Statistical Methods of Time-Series for Estimating and Forecasting the Wheat Series in Yemen (Production and Import). *American Journal of Applied Mathematics*. 2016, vol. 4, no. 3, pp. 124-131. ISSN 2330-0043. e-ISSN 2330-006X. Available from: doi: 10.11648/j.ajam.20160403.12.

4. Allisy-Roberts, P., Williams, J. *Farr's Physics for Medical Imaging*. 2nd edition. Saunders Ltd, 2008. 216 p. ISBN 978-0-7020-2844-1. Available from: doi: 10.1016/C2009-0-34335-4.

5. Asoka, A., Mishra, V. Prediction of Vegetation Anomalies to Improve Food Security and Water Management in India. *Geophysical Research Letters*. 2015, vol. 42, no. 13, pp. 5290-5298. ISSN 0094-8276. e-ISSN 1944-8007. Available from: doi: 10.1002/2015GL063991.

6. Atsalakis, G. S., Skiadas, C. H., Nezis, D. Forecasting Chaotic Time Series by a Neural Network. In: *Proceedings of the 1st Chaotic Modeling and Simulation International Conference, June 3-6, 2008, Chania, Greece*. Chania: CRC Press, 2008, pp. 77-82.

7. Badamasi, M. M., Yelwa, S. A., AbdulRahim, M. A., Noma, S. S. NDVI Threshold Classification and Change Detection of Vegetation Cover at the Falgore Game Reserve in Kano State, Nigeria. *Sokoto Journal of the Social Sciences*. 2015, vol. 2, no. 2, pp. 174-194. ISSN 1595-2738.

8. Bell, A. A., Seiler, C., Kaftan, J. N., Aach, T. Noise in High Dynamic Range Imaging. In: *Proceedings of 15th IEEE International Conference on Image Processing*, October 12-15, 2008, San Diego, USA. IEEE, 2008, pp. 561-564. ISBN 978-1-4244-1765-0. e-ISBN 978-1-4244-1764-3. Available from: doi: 10.1109/ICIP.2008.4711816.

9. Bortolot, Z. J., Wynne, R. H. Estimating Forest Biomass Using Small Footprint LiDAR Data: An Individual Tree-Based Approach That Incorporates Training Data. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2005, vol. 59, no. 6, pp. 342-360. ISSN 0924-2716. Available from:

doi: 10.1016/j.isprsjprs.2005.07.001.

10. Bronson, R., Costa, G. B. *Linear Algebra: An Introduction*, 2nd Edition. USA: Academic Press, 2007. 520 p. ISBN 978-0-12-088784-2.

11. Chakrabarti, G., Sen, C. *Anatomy of Global Stock Market Crashes*. Springer India, 2012. 62 p. ISBN 978-81-322-0462-6. e-ISBN 978-81-322-0463-3. Available from: doi: 10.1007/978-81-322-0463-3.

12. Choi, G., Oh, H. S., Kim, D. Enhancement of Variational Mode Decomposition with Missing Values. *Signal Processing*. 2018, vol. 142, pp. 75-86. ISSN 0165-1684. Available from: doi: 10.1016/j.sigpro.2017.07.007.

13. Clements, N., Sarkar, S., Wei, W. Multiplicative Spatio-Temporal Models for Remotely Sensed Normalized Difference Vegetation Index Data. *Journal of International Energy Policy*. 2014, vol. 3, no. 1, pp. 1-14. ISSN 2165-252X. e-ISSN 2165-2538. Available from: doi: 10.19030/jiep.v3i1.8941.

14. Coulibaly, P., Baldwin, C. K. Nonstationary Hydrological Time Series Forecasting Using Nonlinear Dynamic Methods. *Journal of Hydrology*. 2005, vol. 307, no. 1, pp. 164-174. ISSN 0022-1694. Available from: doi: 10.1016/j.jhydrol.2004.10.008.

15. Crone, S. F. Prediction of White Noise Time Series Using Artificial Neural Networks and Asymmetric Cost Functions. In: *Proceedings of the International Joint Conference on Neural Networks*, July 20-24, 2003, Potland, USA. IEEE, 2003, 3582 p. ISBN 0-7803-7898-9. Available from: doi: 10.1109/IJCNN.2003.1223950.

16. Draper, N. R., Smith, H. *Applied Regression Analysis*. 3rd edition. Hoboken, New Jersey: Wiley, 1998. 736 p. ISBN 978-0-471-17082-2. Available from: doi: 10.1002/9781118625590.ch00.

17. Du, R., Yang, H. An Improved Weighted Moving Average Methods Based on Transferring Weights for an Analytical Process Data. *The Open Automation and Control Systems Journal*. 2014, vol. 6, pp. 1886-1890. ISSN 1874-4443. Available from: doi: 10.2174/1874444301406011886.

18. El-Sharkh, M. Y., Rahman, M. A. Forecasting Electricity Demand Using Dynamic Artificial Neural Network Model. In: *Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management, July 3-6, 2012, Istanbul, Turkey*. 2012, pp. 1691-1694. ISBN 978-0-9855497-0-1.

19. Elman, J. L. Finding Structure in Time. *Cognitive Science*. 1990, vol. 14, no. 2,

pp. 179-211. ISSN 0364-0213. e-ISSN 1551-6709. Available from: doi: 10.1016/0364-0213(90)90002-E.

20. Elsner, J. B. Predicting Time Series Using a Neural Network as a Method of Distinguishing Chaos from Noise. *Journal of Physics A: Mathematical and General*. 1999, vol. 25, no. 4, pp. 843-850. ISSN 0305-4470. Available from: doi:10.1088/0305-4470/25/4/021.

21. Fan, J., Yao, Q. *Nonlinear Time Series: Nonparametric and Parametric Methods*. New York: Springer, 2003. 552 p. ISBN 978-0-387-26142-3. e-ISBN 978-0-387-69395-8. Available from: doi: 10.1007/978-0-387-69395-8.

22. Fattahi, S., Ravandi, S. A. H., Taheri, S. M. Two-Way Prediction of Cotton Yarn Properties and Fiber Properties using Multivariate Multiple Regression. *Journal of the Textile Institute*. 2011, vol. 102, no. 10, pp. 849-856. ISSN 0040-5000. Available from: doi: 10.1080/00405000.2010.524362.

23. Fernandez, M., Caballero, J., Fernandez, L., Sarai, A. Genetic Algorithm Optimization in Drug Design QSAR: Bayesian-regularized Genetic Neural Networks (BRGNN) and Genetic Algorithm-optimized Support Vectors Machines (GA-SVM). *Molecular Diversity.* 2011, vol. 15, no. 1, pp. 269-289. ISSN 1381-1991. e-ISSN 1573-501X. Available from: doi: 10.1007/s11030-010-9234-9.

24. Fernández-Mansoa, A., Quintanob, C., Fernández-Mansoa, O. Forecast of NDVI in Coniferous Areas using Temporal ARIMA Analysis and Climatic Data at a Regional Scale. *International Journal of Remote Sensing.* 2011, vol.32, no. 6, pp. 1595-1617. ISSN 0143-1161. e-ISSN 1366-5901. Available from: doi: 10.1080/01431160903586765.

25. Fletcher, T., Redpath, F., Alessandro, J. D. Machine Learning in FX Carry Basket Prediction. In: *Proceedings of the World Congress on Engineering, vol. II, July 1-3, 2009, London, United Kingdom*. Newswood Limited, 2009, pp. 1371-1375. ISBN 978-988-18210-1-0.

26. Foresti, L., Tuia, D., Timonin, V., Kanevski, M. Time Series Input Selection using Multiple Kernel Learning. In: *Proceedings of the 18th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2010), April 28-30, 2010, Bruges, Belgium*. D-side Publishing, 2010, pp. 123-128. ISBN 2-930307-10-2.

27. Gallaher, D., Campbell, G. G., Meier, W., Moses, J., Wingo, D. The Process of Bringing Dark Data to Light: The Rescue of the Early Nimbus Satellite Data. *GeoResJ*. 2015, vol. 6, pp. 124-134. ISSN 2214-2428. Available from:

doi: 10.1016/j.grj.2015.02.013.

28. Geoimage. *Satellite Overview* [online]. [viewed 2 March 2019]. Available from: https://www.geoimage.com.au/.

29. Grinchenko, N. N., Baranchikov, A. I., Tokarev, A. V. Terrain Objects Edge Detection in Noisy GPS Images. In: *ITM Web of Conferences: 6th Seminar on Industrial Control Systems: Analysis, Modeling and Computation, February 25-26, 2016, Moscow, Russia*. EDP Sciences, 2016, article no. 03004. e-ISSN 2271-2097. Available from: doi: 10.1051/itmconf/20160603004.

30. Haykal, V., Cardot, H., Ragot, N. A Combination of Variational Mode Decomposition with Neural Networks on Household Electricity Consumption Forecast. In: *Proceedings of International Work-Conference on Time Series (ITISE 2017), September 18-20, 2017, Granada, Spain.* Granada: Godel Impresiones Digitales S. L., 2017, pp. 740-752.

31. Hazledine, S., Sun, J., Wysham, D., Downie, J. A., Oldroyd, G. E. D., Morris, R. J. Nonlinear Time Series Analysis of Nodulation Factor Induced Calcium Oscillations: Evidence for Deterministic Chaos? *PLoS ONE*. 2009, vol. 4, no. 8, p. e6637. ISSN 1932-6203. Available from: doi: 10.1371/journal.pone.0006637.

32. Hossain, A., Nasser, M. Recurrent Support and Relevance Vector Machines Based Model with Application to Forecasting Volatility of Financial Returns. *Journal of Intelligent Learning Systems and Applications*. 2011, vol. 3, pp. 230-241. ISSN 2150-8402. e-ISSN 2150-8410. Available from: doi: 10.4236/jilsa.2011.34026.

33. Huang, D., Wu, Z. Forecasting Outpatient Visits using Empirical Mode Decomposition Coupled with Back-Propagation Artificial Neural Networks Optimized by Particle Swarm Optimization. *PLOS One*. 2017, vol. 12, no. 2, p. e0172539. ISSN 1932-6203. Available from: doi: 10.1371/journal.pone.0172539.

34. Huang, N., Yuan, C., Cai, G., Xing, E. Hybrid Short Term Wind Speed Forecasting Using Variational Mode Decomposition and a Weighted Regularized Extreme Learning Machine. *Energies*. 2016, vol. 9, no. 12, pp. 1-19. ISSN 1996-1073. Available from: doi: 10.3390/en9120989.

35. Huang, S. C., Hsieh, C. H. Wavelet-based Relevance Vector Regression Model Coupled with Phase Space Reconstruction for Exchange Rate Forecasting. *International Journal of Innovative Computing, Information and Control*. 2012, vol. 8, no. 3, pp. 1917-1930. ISSN 1349-4198.

36. Huete, A. R. *Environmental Monitoring and Characterization*. Academic Press, 2004. 410 p. ISBN 978-0-12-064477-3. Available from: doi: 10.1016/B978-0-12-064477-3.X5000-0.

37. Jensen, J. R. *Introductory Digital Image Processing: A Remote Sensing Perspective.* 3rd edition. Upper Saddle River, N.J.: Pearson/Prentice Hall, 2005. 526 p. ISBN 9780131453616.

38. Ji, L., Peters, A. J. Forecasting Vegetation Greenness with Satellite and Climate Data. *IEEE Geoscience and Remote Sensing Letters*. 2004, vol. 1, no. 1, pp. 3-6. ISSN 1545-598X. Available from: doi: 10.1109/LGRS.2003.821264.

39. Joo, J. M. Diversity and Temporality of Chaotic Events. *Industrial Data*. 2016, vol. 19, no. 1, pp. 125-130. ISSN 1560-9146. e-ISSN 1810-9993. Available from: doi: 10.15381/idata.v19i1.12545.

40. Kang, L., Di, L., Deng, M. Forecasting Vegetation Index Based on Vegetation-Meteorological Factor Interactions with Artificial Neural Network. In: *Proceedings of the Fifth International Conference on Agro-Geoinformatics (Agro-Geoinformatics 2016), July 18-20, 2016, Tianjin, China*. Tianjin: IEEE, 2016, pp. 1-6. ISBN 978-1-5090-2351-6. e-ISBN 978-1-5090-2350-9. Available from: doi: 10.1109/Agro-Geoinformatics.2016.7577673.

41. Kantz, H., Holstein, D., Ragwitz, M., Vitanov, N. K. Markov Chain Model for Turbulent Wind Speed Data. *Physica A: Statistical Mechanics and its Applications*. 2004, vol. 342, no. 1-2, pp. 315-321. ISSN 0378-4371. Available from: doi: 10.1016/S0378-4371(04)00488-1.

42. Karaca, E., Durmaz, B., Aktug, H., Yildiz, T., Guducu, C., Irgi, M., Gulcihan, M., Koksal, C., Ozkinay, F., Gunduz, C., Cogulu, O. Erratum to: The Genotoxic Effect of Radiofrequency Waves on Mouse Brain. *Journal of Neuro-Oncology*. 2012, vol. 107, no. 3, pp. 665-671. ISSN 1573-7373. e-ISSN 0167-594X. Available from: doi: 10.1007/s11060-012-0850-3.

43. Kay, S. M. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, New Jersey, USA: Prentice-Hall, 1993. 608 p. ISBN 0-13-345711-7.

44. Kayri, M. Predictive Abilities of Bayesian Regularization and Levenberg-Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. *Mathematical and Computational Applications*. 2016, vol. 21, no. 20, pp. 1-11. ISSN 1300-686X. e-ISSN 2297-8747. Available from: doi: 10.3390/mca21020020.

45. Khan, M. J., Yousaf, A., Khurshid, K., Abbas, A., Shafait, F. Automated Forgery

Detection in Multispectral Document Images Using Fuzzy Clustering. In: *Proceedings of 13th IAPR International Workshop on Document Analysis Systems (DAS), April 24-27, 2018, Vienna, Austria*. IEEE, 2018, pp. 393-398. ISBN 978-1-5386-3346-5. e-ISBN 978-1-5386-3347-2. Available from: doi: 10.1109/DAS.2018.26.

46. Khashei, M., Bijari, M. An Artificial Neural Network (p, d, q) Model for Timeseries Forecasting. *Expert Systems with Applications*. 2010, vol. 37, no. 1, pp. 479-489. ISSN 0957-4174. Available from: doi: 10.1016/j.eswa.2009.05.044.

47. Kim, T. Y., Oh, K. J., Kim, C., Do, J. D. Artificial Neural Networks for Non-stationary Time Series. *Neurocomputin*. 2004, vol. 61, pp. 439-447. ISSN 0925-2312. Available from: doi: 10.1016/j.neucom.2004.04.002.

48. Klapetek, P. *Quantitative Data Processing in Scanning Probe Microscopy. SPM Applications for Nanometrology*. 2nd edition. Elsevier, 2018. 416 p. ISBN 978-0-12-813347-7. e-ISBN 978-0-12-813348-4.

49. Klikova, B., Raidl, A. Reconstruction of Phase Space of Dynamical Systems Using Method of Time Delay. In: *WDS'11 Proceedings of Contributed Papers, Part III, May 31-June 3, 2011, Prague, Czech Republic*. Prague: Matfyzpress, 2011, pp. 83-87. ISBN 978-80-7378-186-6.

50. Kochura, Y. P., Stirenko, S., Alienin, O., Novotarskiy, M, Gordienko, Y. G. Performance Analysis of Open Source Machine Learning Frameworks for Various Parameters in Single-Threaded and Multi-Threaded Modes. In: *Proceedings of CSIT 2017: Advances in Intelligent Systems and Computing II, September 5-8, 2017, Lviv, Ukraine*. Springer, Cham, 2017, pp 243-256. ISBN 978-3-319-70580-4. e-ISBN 978-3-319-70581-1. Available from: doi: 10.1007/978-3-319-70581-1_17.

51. Lahmiri, S. A Variational Mode Decomposition Approach for Analysis and Forecasting of Economic and Financial Time Series. *Expert Systems with Applications*. 2016, vol. 55, pp. 268-273. ISSN 0957-4174. Available from: doi: 10.1016/j.eswa.2016.02.025.

52. Lau, K. T., Guo1, W., Kiernan, B., Slater, C., Diamond, D. Non-Linear Carbon Dioxide Determination using Infrared Gas Sensors and Neural Networks with Bayesian Regularization. *Sensors and Actuators B: Chemical*. 2009, vol. 136, no. 1, pp. 242-247. ISSN 0925-4005. Available from: doi: 10.1016/j.snb.2008.11.030.

53. Leine, R., Wouw, N. V. D. *Stability and Convergence of Mechanical Systems*

with Unilateral Constraints. Berlin: Springer, 2008. 236 p. ISBN 978-3-540-76974-3. e-ISBN 978-3-540-76975-0. Available from: doi: 10.1007/978-3-540-76975-0.

54. Leontitsis, A. *Chaotic Systems Toolbox* [online]. [viewed 2 March 2019]. Available from: https://se.mathworks.com/*MATLAB*central/fileexchange/1597-chaotic-systems-toolbox.

55. Li, Z. C., Chiend, C. S., Huang, H. T. Effective Condition Number for Finite Difference Method. *Journal of Computational and Applied Mathematics*. 2007, vol. 198, no. 1, pp. 208-235. ISSN 0377-0427. Available from: doi: 10.1016/j.cam.2005.11.037.

56. Li, Y., Voos, H., Darouach, M., Hua, C. An Application of Linear algebra Theory in Networked Control Systems: Stochastic Cyber-Attacks Detection Approach. *IMA Journal of Mathematical Control and Information*. 2016, vol. 33, no. 4, pp. 1081-1102. ISSN 0265-0754. e-ISSN 1471-6887. Available from: doi: 10.1093/imamci/dnv026.

57. Lindgren, A. C., Johnson, M. T., Povinelli, R. J. Speech Recognition using Reconstructed Phase Space Features. In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03), April 6-10, 2003, Hong Kong, China*. IEEE, 2003, pp. 60-63. ISBN 0-7803-7663-3. Available from: doi: 10.1109/ICASSP.2003.1198716.

58. Liu, Z. Chaotic Time Series Analysis. *Mathematical Problems in Engineering*. 2010, vol. 2010, Article ID 720190, 31 p. ISSN 1024-123X. e-ISSN 1563-5147. Available from: doi: 10.1155/2010/720190.

59. Maaten, L. V. D. *MATLAB Toolbox for Dimensionality Reduction*. [online]. [viewed 2 March 2019]. Available from: https://lvdmaaten.github.io/drtoolbox/.

60. Manobavan, M., Lucas, N. S., Boyd, D. S., Petfor, N. Forecasting the Interannual Trends in Terrestrial Vegetation Dynamics using Time Series Modelling Techniques. In: *ForestSAT Symposium, August 5-9, 2002, Heriot Watt University, Edinburgh, United Kingdom*. London: Kingston University Publishing, 2002, pp. 1-7.

61. Marj, A. F. Agricultural Drought Forecasting using Satellite Images, Climate Indices and Artificial Neural Network. *International Journal of Remote Sensing*. 2011, vol. 32, no. 24, pp. 9707-9719. ISSN 0143-1161. e-ISSN 1366-5901. Available from: doi: 10.1080/01431161.2011.575896.

62. Mitrea, C. A., Lee, C. K. M., Wu, Z. A Comparison between Neural Networks and Traditional Forecasting Methods: A Case Study. *International Journal of*

Engineering Business Management. 2009, vol. 1, no. 2, pp. 19-24. ISSN 1847-9790. Available from: doi: 10.5772/6777.

63. Nai, W., Liu, L., Wang, S., Dong, D. An EMD-SARIMA-Based Modeling Approach for Air Traffic Forecasting. *Algorithms*. 2017, vol. 10, no. 4, article no. 139. ISSN 1999-4893. Available from: doi: 10.3390/a10040139.

64. Naik, J., Dash, S., Dash, P. K., Bisoi, R. Short Term Wind Power Forecasting using Hybrid Variational Mode Decomposition and Multi-Kernel Regularized Pseudo Inverse Neural Network. *Renewable Energy*. 2018, vol. 118, pp. 180-212. ISSN 0960-1481. Available from: doi: 10.1016/j.renene.2017.10.111.

65. NASA. *Landsat Science* [online]. [viewed 2 March 2019]. Available from: https://landsat.gsfc.nasa.gov.

66. NASA. *LP DAAC. Land Process Distributed Active Archive Center* [online]. [viewed 2 March 2019]. Available from: https://lpdaac.usgs.gov/.

67. NASA. *MODIS. Moderate Resolution Imaging Spectroradiometer Data* [online]. [viewed 2 March 2019]. Available from: https://modis.gsfc.nasa.gov/.

68. Neumann, A., Krawczyk, H., Borg, E., Fichtelmann, B. Towards Operational Monitoring of the Baltic Sea by Remote Sensing. In: *Baltcoast 2004 - Managing the Baltic Sea, April 26-28, 2004, Rostock-Warnemünde, Germany*. Rostock-Warnemünde: EUCC, 2004, pp. 211-218. ISSN 0928-2734.

69. Ng'andwe, P., Mwitwa, J., Muimba-Kankolongo, A. *Forest Policy, Economics, and Markets in Zambia*. 1st edition. Academic Press, 2015. 186 p. ISBN 978-0-12-804090-4. e-ISBN 9780128041222. Available from: doi: 10.1016/C2015-0-00973-X.

70. Niaki, S. T. A., Hoseinzade, S. Forecasting S&P Index using Artificial Neural Networks and Design of Experiments. *Journal of Industrial Engineering International*. 2013, vol. 9, no. 1. ISSN 1735-5702. e-ISSN 2251-712X. Available from: doi: 10.1186/2251-712X-9-1.

71. Nordin, F. H., Nagi, F. H., Abidin, A. A. Z. Comparison Study of Computational Parameter Values between LRN and NARX in Identifying Nonlinear Systems. *Turkish Journal of Electrical Engineering and Computer Sciences*. 2013, vol. 21, no. 4, pp. 1151-1165. ISSN 1300-0632. e-ISSN 1303-6203. Available from: doi: 10.3906/elk-1107-12.

72. Oyafuso, M., Carvalho, F. C., Takeshita, T., de Souza, A. L. R., Araújo, D., Merino, V., Gremiao, M. P., Chorilli, M. Development and In Vitro Evaluation of Lyotropic Liquid Crystals for the Controlled Release of Dexamethasone.

Polymers. 2017, vol. 9, no. 8, pp. 330-346. ISSN 2073-4360. Available from: doi: 10.3390/polym9080330.

73. Panigrahi, N. *Computing in Geographic Information Systems*. CRC Press, 2014. 303 p. ISBN 9781482223149. Available from: doi: 10.1201/b17147.

74. Parikh, K. S., Shah, T. P. Support Vector Machine - A Large Margin Classifier to Diagnose Skin Illnesses. *Procedia Technology*. 2016, vol. 23, pp. 369-375. ISSN 2212-0173. Available from: doi: 10.1016/j.protcy.2016.03.039.

75. Plessis, L. D., Xu, R., Damelin, S., Sears, M., Wunsch, D. C. Reducing Dimensionality of Hyperspectral Data with Diffusion Maps and Clustering with K-means and Fuzzy ART. *International Journal of Systems, Control and Communications.* 2011, vol. 3, no. 3, pp. 232-251. ISSN 1755-9340. Available from: doi: 10.1504/IJSCC.2011.042430.

76. Ram, R., Patra, S., Mohanty, M. N. Application of Variational Mode Decomposition on Speech Enhancement. In: *Proceedings of the Second International Conference on Research in Intelligent and Computing in Engineering, March 24-26, 2017, Gopeshwar, Uttrakhand, India*. Warsaw: Polskie Towarzystwo Informatyczne, 2017, pp. 293-296. ISSN 2300-5963. e-ISBN 978-83-65750-05-1. Available from: doi: 10.15439/2017R27.

77. Ribeiro, G. H. T., Neto, P. S. G. de M., Cavalcanti, G. D. C., Tsang, I. R. Lag Selection for Time Series Forecasting using Particle Swarm Optimization. In: *The 2011 International Joint Conference on Neural Networks (IJCNN), July 31-August 5, 2011, San Jose, CA, USA*. IEEE, 2011, pp. 2437-2444. ISBN 978-1-4244-9635-8. e-ISBN 978-1-4244-9637-2. Available from: doi: 10.1109/IJCNN.2011.6033535.

78. Sahebjalal, E., Dashtekian, K. Analysis of Land Use-Land Covers Changes using Normalized Difference Vegetation Index (NDVI) Differencing and Classification Methods. *African Journal of Agricultural Research*. 2013, vol. 8, no. 37, pp. 4614-4622. ISSN 1991-637X. Available from: doi: 10.5897/AJAR11.1825.

79. Saleh, J. M., Hoyle, B. S. Improved Neural Network Performance Using Principal Component Analysis on *MATLAB. International Journal of The Computer, the Internet and Management*. 2008, vol. 16, no. 2, pp. 1-8. ISSN 0858-7027.

80. Sallehuddin, R., Shamsuddin, S. M. H., Hashim, S. Z. M., Abraham, A. Forecasting Time Series Data Using Hybrid Grey Relational Artificial Neural Network and Auto Regressive Integrated Moving Average Model. *Neural*

Network World. 2007, vol. 17, no. 6, pp. 573-605. ISSN 1210-0552.

81. Sallehudin, R., Shamsuddin, S. M. H., Hashim, S. Z. M., Abraham, A. Hybrid Grey Relational Artificial Neural Network and Auto Regressive Integrated Moving Average Model for Forecasting Time-Series Data. *Applied Artificial Intelligence*. 2009, vol. 23, no. 5, pp. 443-486. ISSN 0883-9514. e-ISSN 1087-6545. Available from: doi: 10.1080/08839510902879384.

82. Satellite Imaging Corporation. *Satellite Sensors* [online]. [viewed 2 March 2019]. Available from: https://www.satimagingcorp.com/.

83. Seelan, S. K., Laguette, S., Casady, G. M., Seielstad, G. Remote Sensing Applications for Precision Agriculture: A Learning Community Approach. *Remote Sensing of Environment*. 2003, vol. 88, no. 1, pp. 157-169. ISSN 0034-4257. Available from: doi: 10.1016/j.rse.2003.04.007.

84. Seo, Y., Kim, S., Singh, V. P. Machine Learning Models Coupled with Variational Mode Decomposition: A New Approach for Modeling Daily Rainfall-Runoff. *Atmosphere*. 2018, vol. 9, no. 7, pp. 251-277. ISSN 2073-4433. Available from: doi: 10.3390/atmos9070251.

85. Shabri, A., Samsudin, R. Daily Crude Oil Price Forecasting Using Hybridizing Wavelet and Artificial Neural Network Model. *Mathematical Problems in Engineering*. 2014, vol. 1, pp. 1-10. ISSN 1563-5147. Available from: doi: 10.1155/2014/201402.

86. Soloviev, V., Saptsin, V., Chabenko, D. Markov Chains Application to The Financial-Economic Time Series Prediction. *Computer Modelling and New Technologies*. 2011, vol. 14, no. 3, pp. 16-20. ISSN 1407-5806. e-ISSN 1407-5814.

87. Stepchenko, A. Normalized Difference Vegetation Index Forecasting using a Regularized Layer Recurrent Neural Network. In: *Proceedings of the 3rd Virtual Multidisciplinary Conference QUAESTI, December 7-11, 2015, Zilina, Slovakia*. Zilina: EDIS-Publishing Institution of the University of Zilina, 2015, pp. 261-266. ISBN 978-80-554-1170-5. ISSN 2453-7144. e-ISSN 1339-5572. Available from: doi: 10.18638/quaesti.2015.3.1.192.

88. Stepchenko, A. NDVI Index Forecasting using a Layer Recurrent Neural Network Coupled with Stepwise Regression and the *PCA*. In: *Proceedings of the 5th Virtual International Conference of Informatics and Management Sciences, March 21-25, 2016, Zilina, Slovakia*. Zilina: EDIS-Publishing Institution of the University of Zilina, 2016, pp. 130-135. ISBN 978-80-554-1196-5. ISSN 1339-9144. e-ISSN 1339-231X. Available from:

doi: 10.18638/ictic.2016.5.1.281.

89. Stepchenko, A. Land Cover Classification Based Forecasting on MODIS Imagery Data Using Artificial Neural Networks. In: *Proceedings of the 11th International Scientific and Practical Conference* "*Environment. Technology. Resources*", *June 15-17, 2017, Rezekne, Latvia*. Rezekne: Rezekne Academy of Technologies, 2017, pp. 159-164. ISSN 1691-5402. e-ISSN 2256-070X. Available from: doi: 10.17770/etr2017vol2.2545.

90. Stepchenko, A., Borisov, A. Methods of Forecasting Based on Artificial Neural Networks. *Information Technology and Management Science*. 2014, vol. 17, pp. 25-31. ISSN 2255-9086. e-ISSN 2255-9094. Available from: doi: 10.1515/itms-2014-0003.

91. Stepchenko, A., Chizhov, J. Applying Markov Chains for *NDVI* Time Series Forecasting of Latvian Regions. *Information Technology and Management Science*. 2015, vol. 18, pp. 57-61. ISSN 2255-9086. e-ISSN 2255-9094. Available from: doi: 10.1515/itms-2015-0009.

92. Stepchenko, A., Chizhov, J. *NDVI* Short-Term Forecasting Using Recurrent Neural Networks. In: *Proceedings of the 10th International Scientific and Practical Conference* "*Environment. Technology. Resources", June 18-20, 2015, Rezekne, Latvia*. Rezekne: Rezeknes Augstskola, 2015. ISSN 1691-5402. e-ISSN 2256-070X. Available from: doi: 10.17770/etr2015vol3.167.

93. Stepchenko, A., Chizhov, J. Markov Chain Modelling for Short-Term NDVI Time Series Forecasting. *Information Technology and Management Science*. 2016, vol. 19, pp. 39-44. ISSN 2255-9086. Available from: doi: 10.1515/itms-2016-0009.

94. Stepchenko, A., Chizhov, J., Aleksejeva, L., Tolujew, J. Nonlinear, Non-stationary and Seasonal Time Series Forecasting Using Different Methods Coupled with Data Preprocessing. *Procedia Computer Science*. 2016, vol. 104, pp. 578-585. ISSN 1877-0509. Available from: doi: 10.1016/j.procs.2017.01.175.

95. Sun, G., Chen, T., Wei, Z., Sun, Y., Zang, H., Chen, S. A Carbon Price Forecasting Model Based on Variational Mode Decomposition and Spiking Neural Networks. *Energies*. 2016, vol. 9, no. 1, pp. 1-16. ISSN 1996-1073. Available from: doi: 10.3390/en9010054.

96. Susac, M. Z., Sarlija, N., Pfeifer, S. Combining PCA Analysis and Artificial Neural Networks in Modelling Entrepreneurial Intentions of Students. *Crotian Operational Research Review*. 2013, vol. 4, no. 1, pp. 306-317.

ISSN 1848-0225. e-ISSN 1848-9931.

97. Šāvelis, Rolands. *Signālu diskretizācijas un atjaunošanas paņēmienu izpēte*. Promocijas darbs. Rīga: [RTU], 2013. 132 lpp.

98. Teal, R., Tubana, B., Girma, K., Freeman, K. W., Arnall, D. B., Walsh, O., Raun, W. R. In-Season Prediction of Corn Grain Yield Potential Using Normalized Difference Vegetation Index. *Agronomy Journal*. 2006, vol. 98, no. 6, pp. 1488-1494. ISSN 0002-1962. Available from: doi: 10.2134/agronj2006.0103.

99. Templ, M., Kowarik, A., Filzmoser, P. Iterative Stepwise Regression Imputation using Standard and Robust Methods. *Computational Statistics & Data Analysis*. 2011, vol. 55, no. 10, pp. 2793-2806. ISSN 0167-9473. Available from: doi: 10.1016/j.csda.2011.04.012.

100. Trefethen, L. N., Bau III, D. *Numerical Linear Algebra*. SIAM, 1997. 361 p. ISBN 9780898713619.

101. Vaidyanathan, S. Analysis, Control, and Synchronization of a 3-D Novel Jerk Chaotic System with Two Quadratic Nonlinearities. *Kyungpook Mathematical Journal*. 2015, vol. 55, no. 3, pp. 563-586. ISSN 1225-6951. Available from: doi: 10.5666/KMJ.2015.55.3.587.

102. Valters, Gatis. *FPGA Implementation of Parametrical Orthogonal Transform-Based Experimental DSP Devices*. Promocijas darbs. Rīga: [RTU], 2012. 180 lpp.

103. Vasermanis E., Šķiltere D., Krasts J. *Prognozēšanas metodes.* Rīga: Latvijas Universitāte, 2004. 121 lpp.

104. Ventspils novada pašvaldība. *Ventspils novada teritorijas plānojums* [tiešsaiste]. [skatīts 2019. g. 2. martā]. Pieejams: http://www.ventspilsnovads.lv/images/stories/Teritorijas%20planojumi/2016/Paskaidrojuma_raksts.pdf.

105. Vilde, A., Ruciņš, Ā., Viesturs, D. *Globālās Pozicionēšanas Tehnoloģijas Lauksaimniecībā*. Jelgava: LLU Lauksaimniecības tehnikas zinātniskais institūts, 2008. 47 lpp. ISBN 978-9984-784-47-2.

106. Vuolo, F., Mattiuzzi, M., Klisch, A., Atzberger, C. Data Service Platform for MODIS NDVI Time Series Pre-Processing at BOKU Vienna: Current Status and Future Perspectives. In: *Proceedings of SPIE - Earth Resources and Environmental Remote Sensing/GIS Applications III, September 25-27, 2012, Edinburgh, United Kingdom.* Vol. 8538. SPIE Press, 2012, p. 85380A. ISSN 0277-786X. Available from: doi: 10.1117/12.974857.

107. Wafi, A. F. *Historical Land Use/Land Cover Classification Using Remote Sensing. A Case Study of the Euphrates River Basin in Syria*. Heidelberg:

Springer, 2013. 204 p. ISBN 978-3-319-00623-9. e-ISBN 978-3-319-00624-6. Available from: doi: 10.1007/978-3-319-00624-6.

108. Wang, Y. *Remote Sensing of Coastal Environments (Remote Sensing Applications Series)*. 1st edition. CRC Press, 2009. 458 p. ISBN 1420094416. e-ISBN 9781420094411.

109. Xanthopoulou, G., Salamanis, A., Kehagias, D., Antoniou, I., Bratsas, C., Tzovaras, D. Forecasting Power Output of Photovoltaic Systems Using Linear, Non-Linear and Enhanced Models. In: *Proceedings of International Work-Conference on Time Series (ITISE 2017), September 18-20, 2017, Granada, Spain*. Vol. 1. Godel Impresiones Digitales S.L., 2017, pp. 129-140. ISBN 978-84-17293-01-7786.

110. Yan, Q., Wang, S., Li, B. Forecasting Uranium Resource Price Prediction by Extreme Learning Machine with Empirical Mode Decomposition and Phase Space Reconstruction. *Discrete Dynamics in Nature and Society*. 2014, vol. 5, pp. 1-10. ISSN 1026-0226. Available from: doi: 10.1155/2014/390579.

111. Yusof, Z. M., Abdullah, S., Soaad, S., Yahaya, S. S. S. Comparing the Performance of Modified Ft Statistic with ANOVA and Kruskal Wallis Test. *Applied Mathematics & Information Sciences*. 2013, vol. 7, no. 2L, pp. 403-408. ISSN 1935-0090. e-ISSN 2325-0399. Available from: doi: 10.12785/amis/072L04.

112. Zhang, G. P. Time Series Forecasting using a Hybrid ARIMA and Neural Network Model. *Neurocomputing*. 2003, vol. 50, pp. 159-175. ISSN 0925-2312. Available from: doi: 10.1016/S0925-2312(01)00702-0.

113. Zhang, G., Patuwo, B. E., Hu, M. Y. Forecasting with Artificial Neural Networks: The State of the Art. *International Journal of Forecasting.* 1998, vol. 14, no. 1, pp. 35-62. ISSN 0169-2070. Available from: doi: 10.1016/S0169-2070(97)00044-7.

114. Zhang, G. P., Patuwo, B. E., Hu, M. Y. A Simulation Study of Artificial Neural Networks for Nonlinear Time-Series Forecasting. *Computers & Operations Research*. 2001, vol. 28, no. 4, pp. 381-396. ISSN 0305-0548. Available from: doi: 10.1016/S0305-0548(99)00123-9.

115. Zhang, H., Liang, J., Chai, Z. Stock Prediction Based on Phase Space Reconstruction and Echo State Networks. *Journal of Algorithms & Computational Technology*. 2013, vol. 7, no. 1, pp. 87-100. ISSN 1748-3018. Available from: doi: 10.1260/1748-3018.7.1.87.

116. Zhang, J. S., Xiao, X. C. Predicting Chaotic Time Series Using Recurrent Neural Network. *Chinese Physics Letters*. 2008, vol. 17. no. 2, pp. 88-90. ISSN 0256-307X. Available from: doi: 10.1088/0256-307X/17/2/004.

117. Zhao, J. H., Dong, Z. Y., Xu, Z. Effective Feature Preprocessing for Time Series Forecasting. In: *Proceedings of the Second international conference on Advanced Data Mining and Applications, ADMA 2006, August 14-16, 2006, Xi'an, China*. Springer, 2006, pp. 769-781. ISBN 3-540-37025-0. Available from: doi: 10.1007/11811305_84.

118. Zosso, D., Dragomiretskiy, K. Variational Mode Decomposition. *IEEE Transactions on Signal Processing.* 2013, vol. 62, no. 3, pp. 531-544. ISSN 1053-587X. e-ISSN 1941-0476. Available from: doi: 10.1109/TSP.2013.2288675.

119. Zosso, D. *Variational Mode Decomposition* [online]. 2013 [viewed 2 March 2019]. Available from: https://se.mathworks.com/matlabcentral/fileexchange/44765-variational-mode-decomposition.

120. GIS-Lab. *NDVI - теория и практика* [tiešsaiste]. [skatīts 2019 g. 2. martā]. Pieejams: http://gis-lab.info/qa/NDVI.html.

121. GIS-Lab. *NDVI - теория и [практика]* [tiešsaiste]. [skatīts 2019. g. 2. martā]. Pieejams: http://gis-lab.info/qa/NDVI2.html.

PIELIKUMI

1. pielikums

NDVI indeksa pielietošanas veidi mežsaimniecības nozarē

Pielietošanas veids	Apraksts	Uzdevums
Kompleksa reģionu bioresursu novērtēšana	Bioresursu kompleksā analīze – lapu koku vai skujkoku mežu ar augstu produktivitāti izdalīšana	Analīze, novērtēšana
Mežu produktivitātes novērtēšana	• Mežsaimniecības uzņēmumu salīdzināšana un novērtēšana; • Mežsaimniecības pasākumu plānošana; • Mežu biomasas un krājas modelēšana un skaitliska programmēšana	Analīze, novērtēšana, prognozēšana
Mežsaimniecības efektivitātes novērtēšana	Veikto mežsaimniecības darbu kvalitātes novērtēšana	Analīze, novērtēšana
Mežu ugunsgrēku izplatības monitorings un mežu atjaunošanas kvalitātes analīze	• Mežu ugunsgrēku izplatības prognozēšana; • Ugunsgrēku seku (izdegušo teritoriju) kartēšana; • Izdegušo platību aprēķināšana; • Mežu atjaunošanas procesu monitorings; • Mežu atjaunošanas ātruma un blīvuma novērtēšana	Prognozēšana, kartēšana, analīze, monitorings
Mežu izciršanas monitorings	• Izcirsto teritoriju noteikšana; • Izcirsto teritoriju platības novērtēšana	Analīze, monitorings
Mežu robežu izmaiņu monitorings un analīze	• Mežu izmaiņu novērtēšana; • Mežu platību kartēšanas veikšana	Monitorings, kartēšana, analīze
Nekoksnes mežu resursu novērtēšana	• Nekoksnes mežu resursu ražības modelēšana un skaitliskā prognozēšana; • Nekoksnes mežu resursu platības, korelācijas un citu parametru novērtēšana	Analīze, modelēšana, prognozēšana

*1. pielikums ir veidots, balstoties uz literatūras avota [121] analīzi.

2. pielikums

NDVI indeksa pielietošanas veidi lauksaimniecības nozarē

Pielietošanas veids	Apraksts	Uzdevums
Kultūraugu sējumu un lauku kartēšana	• Kultūraugu sējumu un lauku kartēšana; • Ar kartogrāfiju vai kadastru aprēķinu saistītajos uzdevumos iegūto digitālo karšu izmantošana	Analīze, kartēšana
Kultūraugu ražības monitorings un prognozēšana	• Dažādu kultūraugu ražības skaitlisko prognožu iegūšana; • Veģetācijas izmaiņu dinamikas novērtēšana; • Lauksaimniecības darbu kvalitātes novērtēšana; • Lauksaimniecības darbu plānošana	Analīze, novērtēšana, prognozēšana,
Meliorācijas sistēmu efektivitātes novērtēšana	• Apūdeņošanas un meliorācijas sistēmu stāvokļa novērtēšana; • Apūdeņošanas un meliorācijas sistēmu avārijas vietu noteikšana	Monitorings, analīze, novērtēšana
Zemes izmantošanas monitorings	• Zemju izmantošanas monitorings; • Pamesto lauku atklāšana; • Lauksaimniecības apgrozījuma efektivitātes novērtēšana; • Aizaugušu teritoriju monitorings; • Erozijas un citu dinamisku procesu monitorings	Monitorings, analīze, novērtēšana

*2. pielikums ir veidots, balstoties uz literatūras avota [121] analīzi.

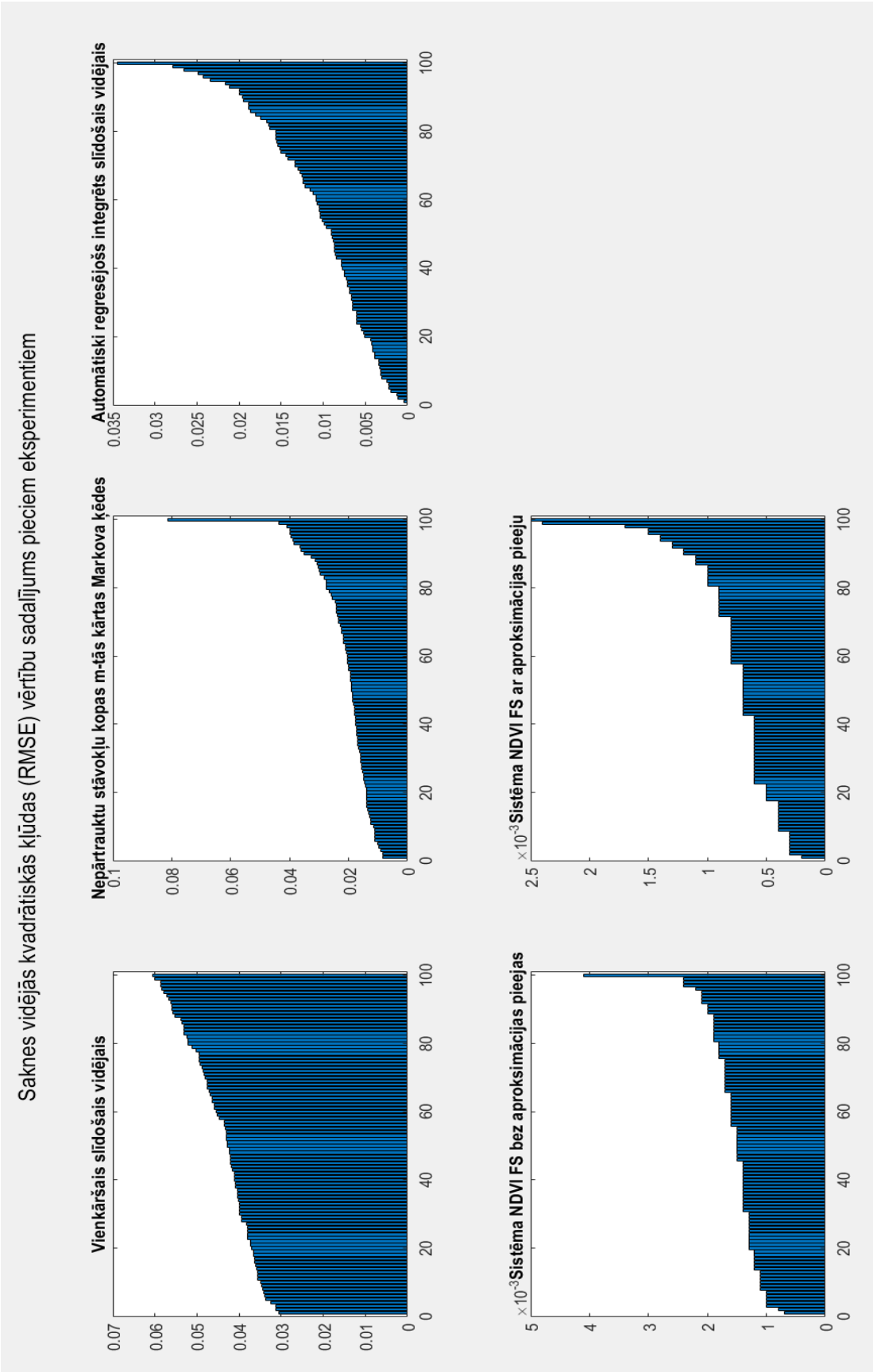
3. pielikums

Zaudējumu funkcijas *RMSE* vērtības pieciem eksperimentiem

Nr.	***SMA***	**Markova ķēdes**	***ARIMA***	***NDVI FS* bez apakšsignāla aproksimācijas pieejas**	***NDVI FS* ar apakšsignāla aproksimācijas pieeju**
1.	0,0537	0,0267	0,0078	0,0011	0,0008
2.	0,0409	0,0149	0,0039	0,0014	0,0011
3.	0,0398	0,0140	0,0065	0,0016	0,0010
4,	0,0364	0,0156	0,0116	0,0011	0,0008
5,	0,0535	0,0307	0,0105	0,0014	0,0010
6,	0,0470	0,0173	0,0199	0,0012	0,0016
7,	0,0465	0,0239	0,0105	0,0018	0,0009
8,	0,0355	0,0185	0,0104	0,0016	0,0009
9,	0,0305	0,0201	0,0012	0,0004	0,0007
10,	0,0465	0,0209	0,0125	0,0015	0,0007
11,	0,0398	0,0200	0,0071	0,0013	0,0010
12,	0,0408	0,0178	0,0039	0,0010	0,0009
13,	0,0571	0,0365	0,0060	0,0015	0,0009
14,	0,0336	0,0131	0,0086	0,0008	0,0007
15,	0,0493	0,0170	0,0200	0,0014	0,0008
16,	0,0326	0,0139	0,0033	0,0009	0,0008
17,	0,0357	0,0180	0,0041	0,0018	0,0010
18,	0,0435	0,0182	0,0032	0,0012	0,0010
19,	0,0431	0,0155	0,0186	0,0007	0,0009
20,	0,0600	0,0389	0,0087	0,0015	0,0010
21,	0,0412	0,0148	0,0069	0,0008	0,0010
22,	0,0365	0,0158	0,0032	0,0008	0,0008
23,	0,0525	0,0296	0,0189	0,0014	0,0008
24,	0,0422	0,0173	0,0096	0,0017	0,0010
25,	0,0370	0,0112	0,0109	0,0012	0,0007
26,	0,0494	0,0147	0,0212	0,0010	0,0010
27,	0,0401	0,0241	0,0030	0,0008	0,0007
28,	0,0459	0,0259	0,0061	0,0012	0,0009
29,	0,0340	0,0169	0,0004	0,0014	0,0009
30,	0,0398	0,0255	0,0055	0,0019	0,0009
31,	0,0313	0,0140	0,0066	0,0017	0,0008
32,	0,0361	0,0085	0,0034	0,0013	0,0009
33,	0,0364	0,0167	0,0164	0,0013	0,0009

Nr.	*SMA*	Markova ķēdes	*ARIMA*	*NDVI FS* bez apakšsignāla aproksimācijas pieejas	*NDVI FS* ar apakšsignāla aproksimācijas pieeju
34,	0,0429	0,0195	0,0071	0,0011	0,0010
35,	0,0400	0,0158	0,0065	0,0011	0,0006
36,	0,0583	0,0396	0,0234	0,0013	0,0010
37.	0,0476	0,0090	0,0243	0,0013	0.0008
38.	0,0521	0,0205	0,0134	0,0009	0.0010
39.	0,0531	0,0277	0,0107	0,0012	0.0009
40,	0,0513	0,0208	0,0266	0,0010	0.0008
41.	0,0381	0,0136	0,0104	0,0011	0.0024
42.	0,0419	0,0195	0,0175	0,0012	0.0009
43.	0,0503	0,0210	0,0090	0,0009	0.0010
44.	0,0380	0,0139	0,0051	0,0013	0.0006
45,	0,0532	0,0399	0,0041	0,0008	0,0009
46,	0,0405	0,0126	0,0025	0,0011	0,0009
47,	0,0521	0,0326	0,0122	0,0012	0,0008
48,	0,0455	0,0243	0,0165	0,0010	0,0006
49,	0,0421	0,0153	0,0087	0,0008	0,0010
50,	0,0561	0,0243	0,0156	0,0018	0,0008
51,	0,0552	0,0351	0,0249	0,0014	0,0010
52,	0,0452	0,0203	0,0195	0,0010	0,0007
53,	0,0412	0,0100	0,0087	0,0007	0,0010
54,	0,0430	0,0112	0,0043	0,0014	0,0010
55,	0,0435	0,0111	0,0134	0,0016	0,0008
56,	0,0379	0,0191	0,0099	0,0010	0,0010
57,	0,0606	0,0300	0,0090	0,0011	0,0010
58,	0,0557	0,0398	0,0196	0,0012	0,0008
59,	0,0382	0,0085	0,0044	0,0013	0,0006
60,	0,0373	0,0190	0,0020	0,0011	0,0006
61,	0,0431	0,0180	0,0060	0,0017	0,0010
62,	0,0495	0,0283	0,0157	0,0011	0,0009
63,	0,0365	0,0178	0,0073	0,0011	0,0009
64,	0,0415	0,0205	0,0034	0,0014	0,0010
65,	0,0480	0,0223	0,0079	0,0015	0,0010
66,	0,0380	0,0140	0,0157	0,0010	0,0009
67,	0,0348	0,0110	0,0085	0,0009	0,0009
68,	0,0314	0,0126	0,0066	0,0016	0,0009

Nr.	*SMA*	Markova ķēdes	*ARIMA*	*NDVI FS* bez apakšsignāla aproksimācijas pieejas	*NDVI FS* ar apakšsignāla aproksimācijas pieeju
69,	0,0420	0,0274	0,0130	0,0011	0,0007
70,	0,0585	0,0315	0,0128	0,0016	0,0012
71,	0,0489	0,0180	0,0345	0,0016	0,0013
72,	0,0494	0,0228	0,0061	0,0007	0,0010
73,	0,0460	0,0225	0,0145	0,0007	0,0009
74,	0,0421	0,0193	0,0065	0,0007	0,0006
75,	0,0342	0,0168	0,0052	0,0012	0,0007
76,	0,0559	0,0276	0,0279	0,0011	0,0010
77,	0,0559	0,0386	0,0167	0,0018	0,0011
78,	0,0395	0,0219	0,0150	0,0011	0,0009
79,	0,0486	0,0216	0,0154	0,0013	0,0012
80,	0,0532	0,0216	0,0152	0,0010	0,0010
81,	0,0422	0,0187	0,0056	0,0011	0,0010
82,	0,0429	0,0159	0,0022	0,0005	0,0009
83,	0,0346	0,0099	0,0109	0,0013	0,0010
84,	0,0404	0,0140	0,0069	0,0016	0,0009
85,	0,0432	0,0233	0,0124	0,0010	0,0010
86,	0,0404	0,0176	0,0155	0,0011	0,0009
87,	0,0586	0,0302	0,0022	0,0021	0,0014
88,	0,0477	0,0246	0,0089	0,0021	0,0006
89,	0,0566	0,0410	0,0075	0,0013	0,0010
90,	0,0359	0,0143	0,0088	0,0009	0,0009
91,	0,0483	0,0161	0,0189	0,0010	0,0010
92,	0,0372	0,0128	0,0101	0,0012	0,0013
93,	0,0477	0,0190	0,0124	0,0010	0,0008
94,	0,0395	0,0438	0,0075	0,0038	0,0023
95,	0,0446	0,0188	0,0216	0,0014	0,0007
96,	0,0578	0,0360	0,0112	0,0014	0,0010
97,	0,0345	0,0116	0,0011	0,0009	0,0009
98,	0,0412	0,0172	0,0142	0,0016	0,0014
99,	0,0469	0,0234	0,0180	0,0015	0,0010
100,	0,0357	0,0814	0,0077	0,0021	0,0005

4. pielikums

Zaudējumu funkcijas *RMSE* vērtību sadalījums