УЎТ631.4

СУНЪИЙ ИНТЕЛЛЕКТ ТЕХНОЛОГИЯЛАРИ ЁРДАМИДА ТУПРОҚ МАЪЛУМОТЛАРИНИ ИНТЕЛЛЕКТУАЛ ТАҲЛИЛГА ТАЙЁРЛАШ

Баҳодиров Зафар Абдувалиевич,
Тупроқшунослик ва агрокимёвий тадқиқотлар институти,
бўлим мудири, e-mail: zafarbahodirov@gmail.com

Аннотация. Мақолада сунъий интеллект технологиялари ёрдамида тупроқ маълумотларини интеллектуал таҳлилга тайёрлаш жараёни ёритилган. Бунда тупроқ маълумотларини интеллектуал таҳлилга тайёрлашда визуализация, корреляция ва стандартлаш каби босқичларни амалга оширишга оид таҳлиллар келтирилган.

Калит сўзлар: тупроқ маълумотлари, сунъий интеллект, интеллектуал таҳлил, визуализация, корреляция, стандартлаш.

Аннотация. В статье описан процесс подготовки почвенных данных для интеллектуального анализа с использованием технологий искусственного интеллекта. При этом представлены такие этапы подготовки почвенных данных для интеллектуального анализа, как визуализация, корреляция и стандартизация.

Ключевые слова: почвенные данные, искусственный интеллект, интеллектуальный анализ, визуализация, корреляция, стандартизация.

Annotation. The article describes the process of preparing soil data for intellectual analysis using artificial intelligence technologies. This analysis introduces steps such as visualization, correlation, and standardization in preparing soil data for mining.

Key words: soil data, artificial intelligence, intellectual analysis, visualization, correlation, standardization.

Кириш. Сўнгги 50 йил ичида дунё аҳолиси 3 миллиарддан 7 миллиардга ошиб, озиқ-овқат маҳсулотларига бўлган юқори талабни яратди. Бирлашган Миллатлар Ташкилотининг Озиқ-овқат ва қишлоқ хўжалиги ташкилоти тахминларига кўра, 2050 йилгача дунё аҳолиси 30 фоизга ошади, демак, озиқ-овқат маҳсулотларини 70 фоизга ошишига эришиш керак. Ерларнинг деградацияга учраши, иқлим ўзгариши ва бозорнинг ўзгариши озиқ-овқат хавфсизлигида муаммоларни юзага келтирмоқда.

Ҳозирги кунда қишлоқ хўжалигида турли замонавий технологиялар, хусусан, рақамли технологиялар қўлланилмоқда, масалан, масофадан туриб зондлаш, булутли ҳисоблаш ва Интернет нарсалар (IoT) кабилардан кенг фойдаланилмоқда. Сунъий интеллект технологиялари кўплаб соҳаларда муваффақиятли қўлланилиб келаётган бўлсада, у қишлоқ хўжалигида яқиндан, манфаатдор томонлар унинг потенциал фойдаларини англай бошлагандангина қўлланила бошланди. Баъзи йирик қишлоқ хўжалиги корпорацияларининг фикрига кўра, сунъий интеллект технологиялари асосида фермерларга маслаҳатлар бериш дунё миқёсида экинлардан олинадиган йиллик глобал даромадни тахминан 20 миллиард АҚШ долларига кўпайтириши мумкин.

Республикада тупроқ тадқиқотлари натижаларини таҳлил қилишда янги технологиялар, хусусан, сунъий интеллект таҳлили усулларидан фойдаланиш тупроқшунослик фанини янги даражага чиқаришда катта аҳамиятга эга. Республикамизда барча соҳаларида инновацион технологияларни тадбиқ этиш, шунингдек, ер ресурсларини бошқаришда автоматлаштирилган геоахборот тизимини қўллаш ушбу тизимини етакчи ўринларга олиб чиқишда бугунги куннинг энг долзарб масаласи ҳисобланади. Ушбу соҳада сунъий интеллект технологияларини қўллаш эса замонавий компьютер дастурлари ва қурилмалари асосида маълум бир обьект ёки ҳудуд тўғрисидаги тўлиқ ахборотларни таҳлил этиш ва маълум йўналишдаги масалаларни ечиш, шунингдек, ўрганилаётган ҳудудлар геоахборот базаси ва моделини ишлаб чиқиш, қишлоқ хўжалиги экинларида кечадиган салбий ҳолатларни олдиндан башоратлаш ва бунинг ҳисобига эса ўсимлик ва ер ресурсларини самарали бошқариш имкониятини яратади.

Бугунги кунда тупроқшунослик соҳасида сунъий интеллект технологияларидан фойдаланиш бўйича кенг кўламли илмий ишлар олиб борилмоқда. Жумладан, МсБратней ва бошқалар томонидан таклиф қилинган рақамли тупроқ хариталаш тизими орқали ёрдамчи ўзгарувчилар ва тупроқ хусусиятларини ўзаро боғлаш учун кўплаб башорат моделлари ишлаб чиқилган ва жорий этилган [1]. Масалан, Минасни ва бошқалар. [2] тупроқ органик моддасини моделлаштириш ва тупроқ органик моддасининг атроф-муҳит ковариативлари билан алоқасини ҳар томонлама таҳлил қилди. Шунингдек, Ҳенгл ва бошқалар [3] кўп ва чизиқли регрессия усуллари орқали тупроқ хоссалари ва ёрдамчи ўзгарувчилар

боғлиқлигини таҳлил қилдилар. Тупроқ хоссаларининг тупроқ профилидаги вертикал тақсимланиши ландшафт миқёсида моделлаштирилган ва башорат қилувчи сифатида сплайнлар усули, сунъий интеллект воситалари ва атроф-муҳит ковариациялари ёрдамида башорат қилинган [4, 5]. Валдер ва бошқалар [6] Германиядаги тупроқ текстураси компонентлари ва гумус миқдорини башорат қилиш учун геостатистик усуллардан фойдаланган. Анъанавий ёндашувлар билан солиштирган ҳолда тупроқ хоссаларини ўрганиш учун умумлаштирилган чизиқли моделлар [7], регрессия дарахти моделлари [8], тасодифий ўрмон [9], кубист [10], вектор регрессиясини қўллаб-қувватлаш [11] ва тасодифий ўрмон регрессияси [11] каби сунъий интеллект алгоритмларидан фойдаланилган.

Тадқиқот объекти ва услублари. Тадқиқот объекти сифатида Сирдарё вилояти Мирзаобод тумани тупроқлари олинган. Тадқиқотларда тупроқларнинг механик таркиби, озиқа моддалар билан таъминланганлик даражаси ва шўрланиш ҳолати ўрганилган. Тупроқ маълумотларини интеллектуал таҳлилга тайёрлашда сунъий интеллект технологиялари соҳасида кенг қўлланилувчи CRSIP-DM методологиясидан фойдаланилди. Маълумотлар базаси маълумотларни таҳлил қилишда Anaconda дастурий таъминотидан фойдаландик. Дастурлаш тили сифатида сунъий интеллект технлогияларида кенг қўлланилуви python дан фойдаланилди. Бунда унинг pandas, numpy, seaborn, scikit-learn кутубхоналари қўлланилди.

Тадқиқот натижалари. Сунъий интеллект технологияларидан фойдаланган ҳолда таҳлиллар олиб бориш учун лойиҳа объекти бўлган Сирдарё вилояти Мирзабод тумани тупроқлари ҳолати бўйича маълумотлар базаси ишлаб чиқилди. Ушбу базада туман тупроқларининг механик таркиби, гумус ва озиқа элементлари билан таъминланганлиги ҳамда шўрланиш ҳолатига оид маълумотлар шакллантирилди.

Сўнгра сунъий интеллект технологияларидан фойдаланган ҳолда таҳлиллар олиб бориш ва интеллектуал моделлар ишлаб чиқиш учун сунъий интеллект технологиялари соҳасида кенг қўлланилувчи CRSIP-DM методологиясидан фойдаланди. Ушбу методология олиб бориладиган барча таҳлиллар кетма-кетлигини ўз ичига олиб, сунъий интеллект технологиялари асосида таҳлилларни аввалидан бошлаб ишлаб чиқилган моделни тақдимотигача бўлган қадамларни белгилаб беради.

Ушбу методологиянинг маълумотларни ўрганиш босқичида ишлаб чиқилган маълумотлар базаси бўйича тегишли маълумотлар таҳлил қилинди. Бунда ишлаб чиқилган маълумотлар базаси сунъий интеллект маълумотлари таҳлили олиб бориш учун бир неча тайёргарлик босқичларида таҳрирланди.

Дастлаб маълумотлар базасидаги устунлар бўйича таҳлил қилиниб, баъзи устунлар бўйича маълумотлар (қаторлар) жуда камлиги аниқланди. Булар тупроқдаги Ca, Mg, Na миқдорлари, 0,01-0,005, 0,005-0,001, <0,001 оралиғидаги заррачалар миқдорлари ҳамда тупроқдаги гипс миқдори устунлари бўлди.

Сунъий интеллект асосидаги таҳлиллар учун бу каби устунлар катта хатоликлар келтириб чиқариши сабабли ушбу устунлар маълумотлар базасидан чиқариб ташланди.

Кейин эса маълумотлар базаси ва унга мос келувчи тупроқ кесмалари ўртасида таҳлиллар олиб борилиб, фақатгина мос келувчи қаторлар олинди. Ортиқча ёки етмаётган қаторлар олиб ташланди.

Ҳосил бўлган маълумотлар базаси маълумотларни таҳлил қилиш усулларидан фойдаланиб ўрганилди. Бунда Anaconda дастурий таъминотидан фойдаланилди.

Бунда дастлаб бизга керакли модуллар чақириб олинди. Яъни pandas, numpy, seaborn, scikit-learn кутубхоналари чақирилди.

```
import pandas as pd
import numpy as np
import sklearn # scikit-learn kutubxonasi
import matplotlib.pyplot as plt
import seaborn as sns
```

Тегишли маълумотлар базаси ўқитиб олинди:

```
filepl=«https://github.com/ageron/handson-ml2/blob/master/datasets/housing/housing.csv?raw=true»
df = pd.read_csv(filepl)
```

Аввалои датасет бўйича қисқача маълумотларни кўриб чиқамиз:

```
df.head()        df.info()        df.describe()
```

```
Ввод [19]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 657 entries, 0 to 656
Data columns (total 17 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   massiv       657 non-null    object
 1   id           657 non-null    int64
 2   QuruqQoldiq  657 non-null    float64
 3   HCO3         657 non-null    float64
 4   Cl           657 non-null    float64
 5   SO4          657 non-null    float64
 6   SalTypeCla   657 non-null    int64
 7   SaldegreeC   657 non-null    int64
 8   Mless025     657 non-null    float64
 9   M025to01     657 non-null    float64
 10  M01to005     657 non-null    float64
 11  M005to001    657 non-null    float64
 12  PhysClay     657 non-null    float64
 13  MechCont_1   657 non-null    int64
 14  Humus        657 non-null    float64
 15  P2O5         657 non-null    float64
 16  K2O          657 non-null    int64
```

```
Ввод [21]: df.describe()
```

Out[21]:

	id	QuruqQoldiq	HCO3	Cl	SO4	SalTypeCla	SaldegreeC	Mless025	M025to01	M01to005	M005to001	PhysClay
count	6.570000e+02	657.000000	657.000000	657.000000	657.000000	657.000000	657.000000	657.000000	657.000000	657.000000	657.000000	657.000000
mean	5.293539e+05	1.390647	0.029385	0.279318	0.548501	2.788432	3.337900	2.993607	0.753577	20.715982	48.105175	27.174429
std	6.934304e+06	1.587570	0.006352	0.495765	0.626840	1.322641	1.196404	2.817772	0.704725	10.583550	10.931362	8.169281
min	1.011000e+03	0.110000	0.018000	0.007000	0.033000	1.000000	1.000000	0.400000	0.100000	1.000000	11.200000	8.100000
25%	1.067800e+04	0.430000	0.024000	0.032000	0.157000	1.000000	2.000000	1.200000	0.300000	13.000000	41.300000	22.200000
50%	1.176700e+04	0.840000	0.027000	0.091000	0.308000	3.000000	3.000000	2.000000	0.500000	20.000000	48.200000	26.300000
75%	1.091920e+05	1.645000	0.033000	0.287000	0.730000	4.000000	5.000000	4.000000	1.000000	26.500000	55.400000	31.600000
max	1.041892e+08	11.130000	0.061000	5.250000	4.833000	4.000000	5.000000	19.600000	4.900000	54.100000	78.400000	69.700000

Train set ва Test set.

Сунъий интеллект асосидаги таҳлиллар учун маълумотларни икки қисмга ажратиб олиш зарур:

- Train set – модел яратиш учун;
- Test set – модел аниқлигини текшириш учун.

Маълумотларнинг 80% train, 20% test учун ажратамиз. Бунинг учун scikit-learn таркибида тайёр train_test_split функциясига мурожаат қиламиз.

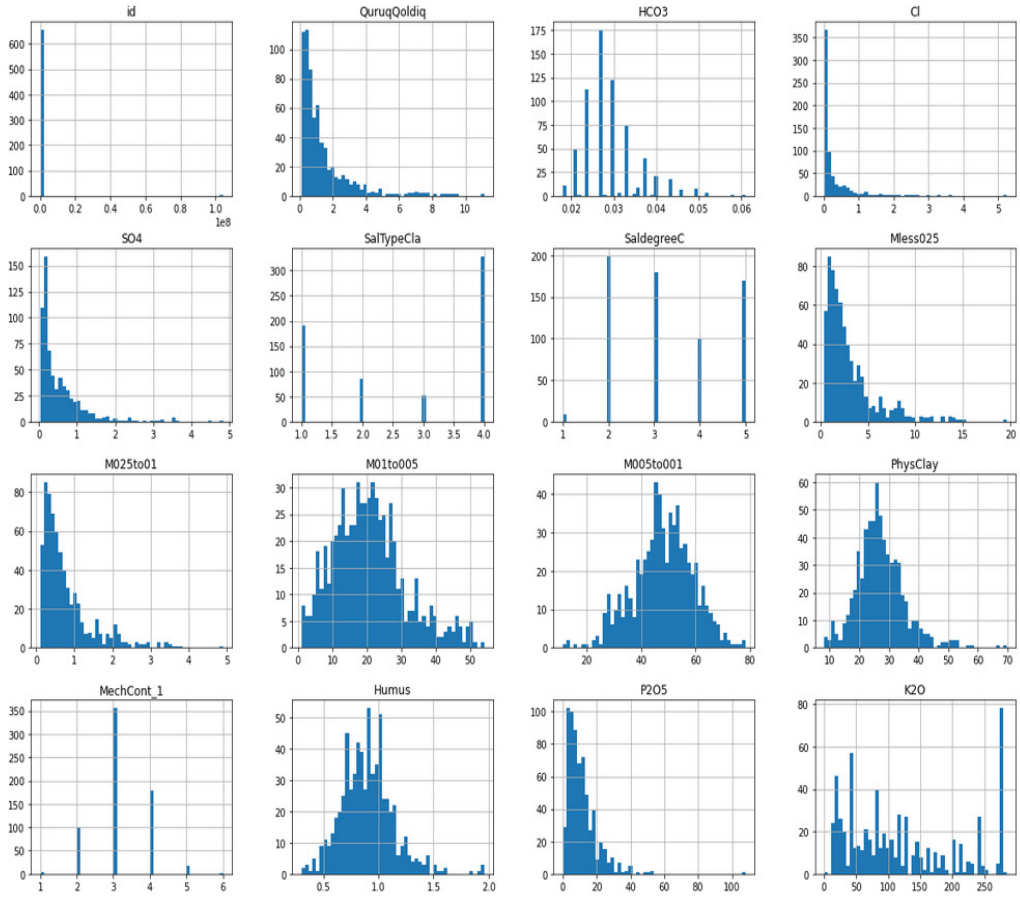
Функцияга параметр сифатида dataset (df), test set ҳажми (0,2 яъни 20%) ва тасодифий сонлар генератори учун қиймат (random_seed) берамиз.

random_seed нинг вазифаси train_test_split функцияни ишга туширганда доим бир хил тасодифий қийматлар олиш. Бу эса, МЛ модель яратиш жараёнида test_set доим яширин қолишини таъминлайди.

```
from sklearn.model_selection import train_test_split
train_set, test_set = train_test_split(df, test_size=0.2, random_state=42)
train_set
```

Визуализация

```
%matplotlib inline
df.hist(bins=50, figsize=(20,15))
plt.show ()
```




```
test_set.shape (132, 17)
```

Маълумотларни таҳлил қилиш:

Келинг энди бевосита маълумотларни таҳлил қилишга ўтайлик. Эсингизда бўлса биз энди train set билан ишлашимиз керак.

Қулайлик учун бу сетимиздан нусха кўчириб оламиз.

tuproq = train_set.copy()

Корреляция

Бизнинг асл мақсадимиз бизга берилган маълумотлар орасида тупроқ ҳолатига (гумус, озиқа моддалар, шўрланиш, механик таркиб) таъсир қилувчи параметрларни топиш. Бунда эса бизга айнан корреляция жуда қўл келади.

tuproq.corrwith(housing['QuruqQoldiq']).sort_values(ascending=False)

```
Ввод [27]: tuproq.corrwith(tuproq['QuruqQoldiq']).sort_values(ascending=False)

Out[27]: QuruqQoldiq    1.000000
         Cl             0.848397
         SO4            0.827392
         SaldegreeC     0.667140
         Mless025       0.193206
         M025to01       0.189146
         K2O            0.132325
         Humus          0.087278
         SalTypeCla     0.070615
         M005to001      0.050913
         P2O5          -0.022825
         id            -0.038921
         HCO3          -0.039214
         M01to005      -0.055402
         PhysClay      -0.076581
         MechCont_1    -0.083689
         dtype: float64
```

Кўришимиз мумкинки, Quruq Qoldiq устуни билан Cl, SO4, SalDegreeC устунлари орасида корреляция нисбатан кучли.

seaborn таркибидаги pairplot функцияси ёрдамида корреляция қийматларини график кўринишида чиқаришимиз ҳам мумкин.

cols = ['Quruq Qoldiq','Cl',' SO4','SaldegreeC']

sns.pairplot(tuproq [cols], height=5)

plt.show()

1. Кейинги босқич маълумотларни интеллектуал таҳлилга тайёрлаш бўлиб, бунда маълумотлар орасидаги йўқ қийматларни таҳрир қилиш, матнли устунларни ўзгартириш ва қийматларни стандартлаш амаллари бажарилди.

Ушбу босқичда тупроқлар ҳолати маълумотлар базасидан йўқ қийматли барча қаторлар чиқариб ташланди. Сўнгра тупроқ шўрланиш типи, шўрланиш даржаси ва механик таркиб тури каби матнли устунларга махсус кодлар берилиб шифрланди. Шунингдек, ушбу босқичда массивлар номлари ҳам кодланди.

Salinity type	**Class**
С	1
Х	2
с/х	3
х/с	4
Salinity degree	**Class**
Шўрланмаган	1
Кам	2
Ўрта	3
Кучли	4
Жуда кучли	5

Mechanical content	**Class**
Қум	1
Қумлоқли	2
Енгил	3
Ўрта	4
Оғир	5
Лой	6

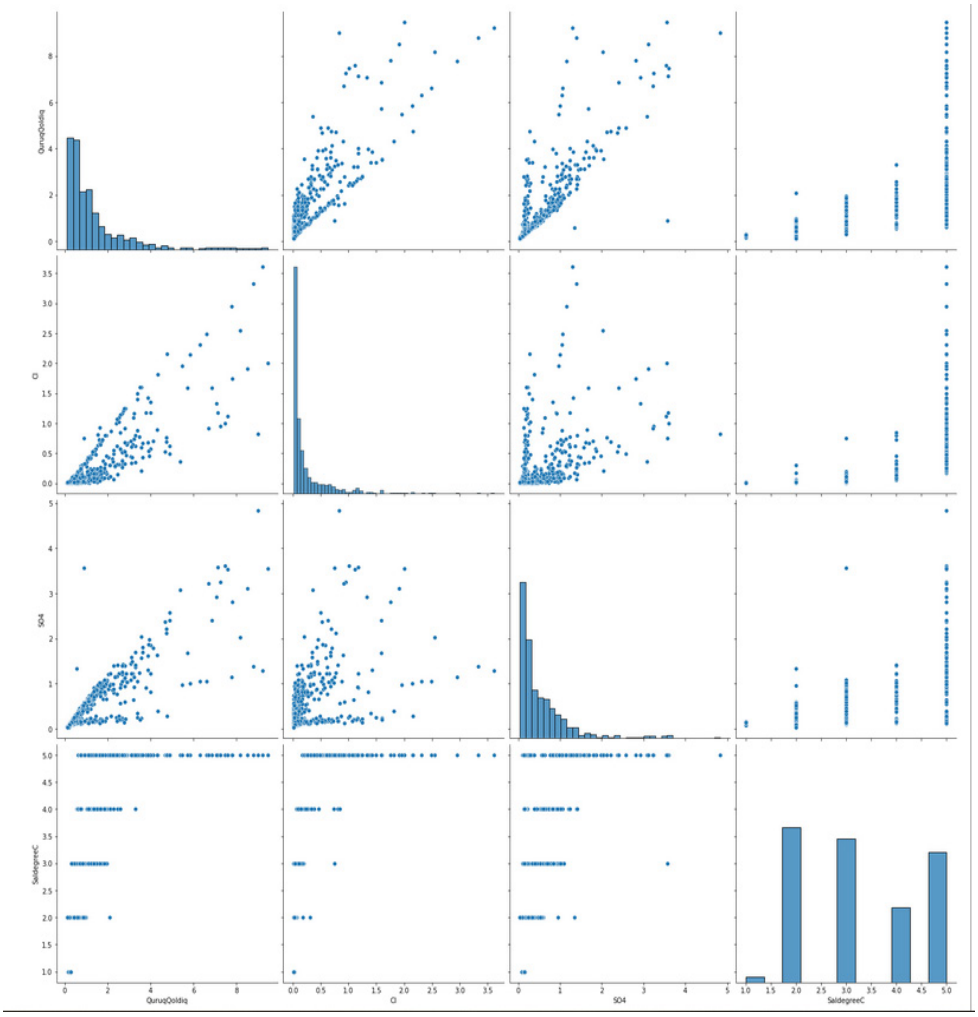
Standard Scaler - стандартлаш

Standard Scaler – сонларни стандартлаш учун бизга берилган устун бўйича ўрта қиймат (μ) ва стандарт

чекланиш (σ) талаб қилинади. Ҳар бир қатордаги қиймат эса қуйдагича ҳисобланади:

```
tuproq_num = housing.drop(«massiv», axis=1)
from sklearn.preprocessing import StandardScaler
standart_scaler = StandardScaler()
standart_scaler.fit_transform(tuproq_num)
```

```
Out[31]: array([[-6.85397815e-02,  1.06822574e+00, -1.31233983e+00, ...,
                  6.21422664e-01, -4.88560148e-01,  1.07733654e-01],
                [-6.88053100e-02,  1.63628064e+00,  1.59409663e+00, ...,
                 -1.25037265e-01, -6.98429624e-01, -1.16805962e+00],
                [-6.86702867e-02, -6.17003801e-01, -8.53428809e-01, ...,
                  9.28788517e-01, -9.08299099e-01, -1.87588862e-01],
                ...,
                [-6.90633515e-02, -4.15028725e-01,  1.13518561e+00, ...,
                 -2.56765488e-01, -5.51796819e-02, -9.31801603e-01],
                [-8.22358630e-02, -4.18184586e-01,  8.29244927e-01, ...,
                  9.45097725e-02,  1.44118968e+00, -1.07355641e+00],
                [-6.97064546e-02, -3.45599793e-01, -3.94517790e-01, ...,
                  1.82328588e-01, -2.78690673e-01,  1.41754808e-03]])
```



Хулоса. Тупроқ ҳолатини сунъий интеллект технологиялари ёрдамида таҳлил қилиш истиқболли бўлиб, ишлаб чиқилган интеллектуал моделлар асосида тупроқ унумдорлик кўрсаткичларини турли аниқликларда башорат қилиш мумкин. Бу эса ўз навбатида тупроқ унумдорлик ҳолатини тезкор аниқлаш ва шу асосда тегишли тадбирларни ишлаб чиқиш имкониятларини яратади.

Фойдаланилган адабиётлар рўйхати:

1.McBratney, A.B., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3, 52
2.Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Chapter one - digital mapping of soil carbon. In: Sparks, D.L. (Ed.), Advances in Agronomy. Academic Press, pp. 1-47
3.Hengl, T., Heuvelink, G.B., Kempen, B., Leenaars, J.G., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., de Jesus, J.M., Tamene, L., 2015. Mapping soil properties of Africa at 250 m resolution: random forests significantly improve current predictions. PLoS One 10, e0125814.
4.Malone, B.P., McBratney, A., Minasny, B., Laslett, G., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. Geoderma 154 (1–2), 138-152.
5.Lacoste, M., Minasny, B., McBratney, A., Michot, D., Viaud, V., Walter, C., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. Geoderma 213-296-311.
6.Wälder, K., Wälder, O., Rinklebe, J., Menz, J., 2008. Estimation of soil properties with geostatiscal methods in floodplains. Arch. Agron. Soil Sci. 54 (3), 275-295.
7.Karunaratne, S., Bishop, T., Baldock, J., Odeh, I., 2014. Catchment scale mapping of measureable soil organic carbon fractions. Geoderma 219-14-23.
8.Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., Malone, B., 2014. Digital mapping of soil salinity in Ardakan region, central Iran. Geoderma 213-15-28.
9.Pahlavan-Rad, M.R., Akbarimoghaddam, A., 2018. Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). Catena 160-275-281.
10.Taghizadeh-Mehrjardi, R., Nabiollahi, K., Kerry, R., 2016. Digital mapping of soil organic carbon at multiple depths using different data mining techniques in Baneh region, Iran. Geoderma 266-98-110.
11.Forkuor, G., Hounkpatin, O.K., Welp, G., Thiel, M., 2017. High resolution mapping of soil properties using remote sensing variables in South-Western Burkina Faso: a comparison of machine learning and multiple linear regression models. PLoS One 12, e 0170478.

УЎТ: 631.41.2

ЧИРЧИҚ-АНГРЕН ОКРУГИ ТОҒ ОСТКИ АДИРЛИКЛАРИДА ТАРҚАЛГАН ТИПИК БЎЗ ТУПРОҚЛАРНИНГ АСОСИЙ КИМЁВИЙ ХОССАЛАРИ

Тошқўзиев Маруф Мансурович,
б.ф.д., профессор, *maruf41@rambler.ru*
Каримов Хаитоли Хурсанович,
таянч докторант, *hayitali.karimov.1985@mail.ru*

Тупроқшунослик ва агрокимёвий тадқиқотлар институти

Аннотация. Мақолада Оҳангарон дарёси ҳавзасида тарқалган қўриқ, лалми ва суғориладиган типик бўз тупроқларда гумус, умумий озиқа моддалари, карбонатлар миқдори тарқалишини уларни шакллани-шига боғлиқ ҳолда тақсимланишига доир янги маълумотлар келтирилган. Ушбу моддаларнинг миқдори ва тарқалиши тупроқларнинг рельеф бўйича жойлашганлиги, табиий ҳамда инсон омиллари таъсири билан боғлиқлиги аниқланган.

Калит сўзлар: қўриқ, лалми, суғориладиган, типик бўз тупроқ, гумус, умумий озиқа моддалари, карбонатлар.

Аннотация. В статье приведены новые сведения по распространению содержания гумуса, общих пи-тательных веществ и карбонатов в целинных, богарных и орошаемых типичных сероземах в бассейне р. Ахангаран в зависимости от условий их формирования. Отмечено, что содержание и распространение этих веществ связано с расположением почв по рельефу, природными и антропогенными факторами.

Ключевые слова: целинный, богарный, орошаемый, типичный серозем, гумус, общие питательные вещества, карбонаты.

Annotation. The article presents new information on the distribution of humus content, total nutrients and carbonates depending on the conditions of their formation in virgin, rainfed and irrigated typical serozem soils