

**Разработка автоматизированной системы совместного анализа  
спутниковых и статистических данных для оценки  
урожайности сельскохозяйственных культур**

Сычков Александр Александрович

Трошко Ксения Анатольевна

# Разработка автоматизированной системы совместного анализа спутниковых и статистических данных для оценки урожайности сельскохозяйственных культур

№1

Мониторинг состояния с/х культур имеет важное значения для оценок их фактического состояния и расчетного объема. Такого рода оценки представляют интерес как для государственных структур, так и для коммерческих организаций. Важное значение имеют методы, позволяющие получить однородные и сопоставимые оценки состояния озимых культур на больших территориях. Для этих целей необходимо использовать данные спутниковых наблюдений.

В основе процедуры лежит анализ сезонной динамики вегетационного индекса NDVI и ее связи с метеорологическими условиями на изучаемых территориях. Временные ряды формируются на конкретных территориях на основе еженедельных композитов спектрометра MODIS. (Спутники Terra и Aqua)

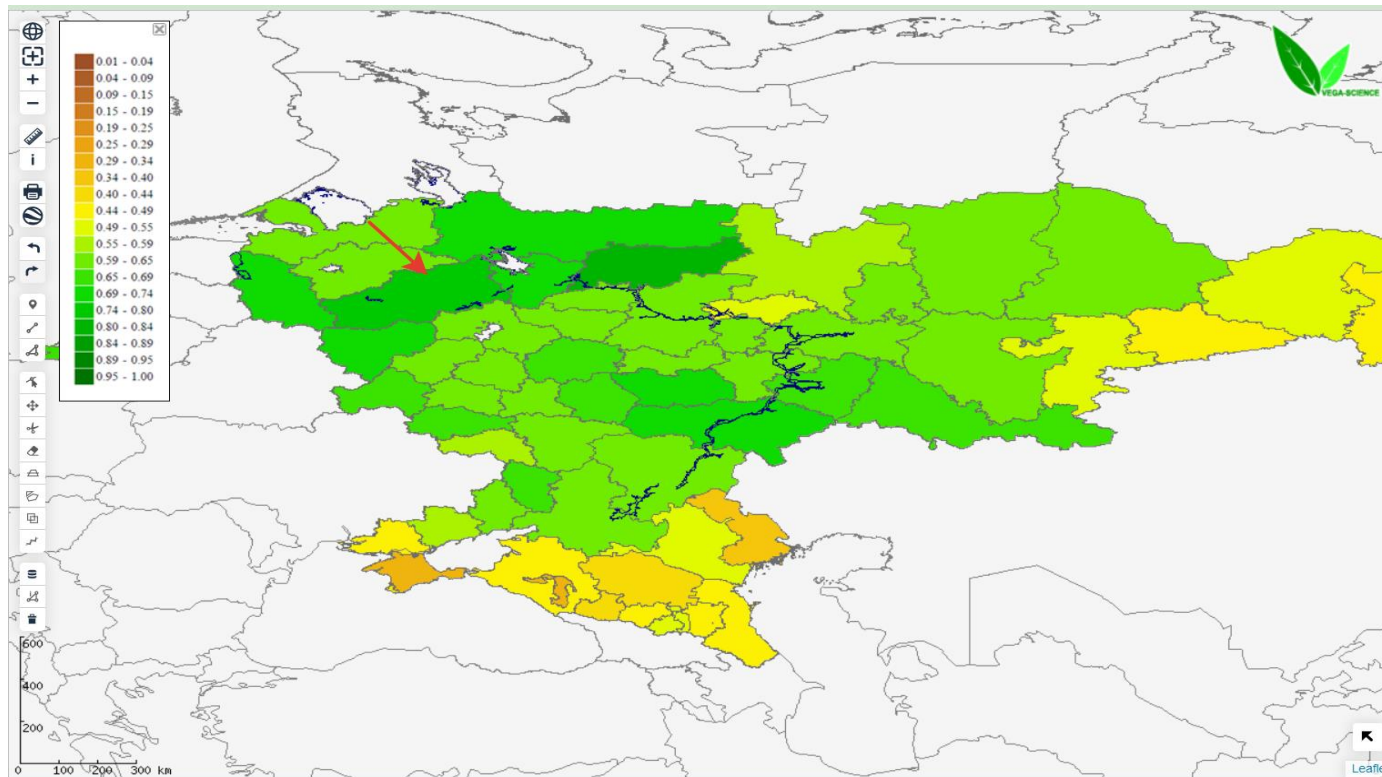
**Цель работы:** Создать автоматизированную систему оценки состояния урожайности на основе спутниковых и статистических данных.

## **Задачи:**

1. Разработать программы для автоматизированного сбора статистической и спутниковой информации
2. Предусмотреть возможность ввода дополнительной информации
3. Разработать программу для автоматизированной обработки информации
4. Автоматизировать формирование отчетов

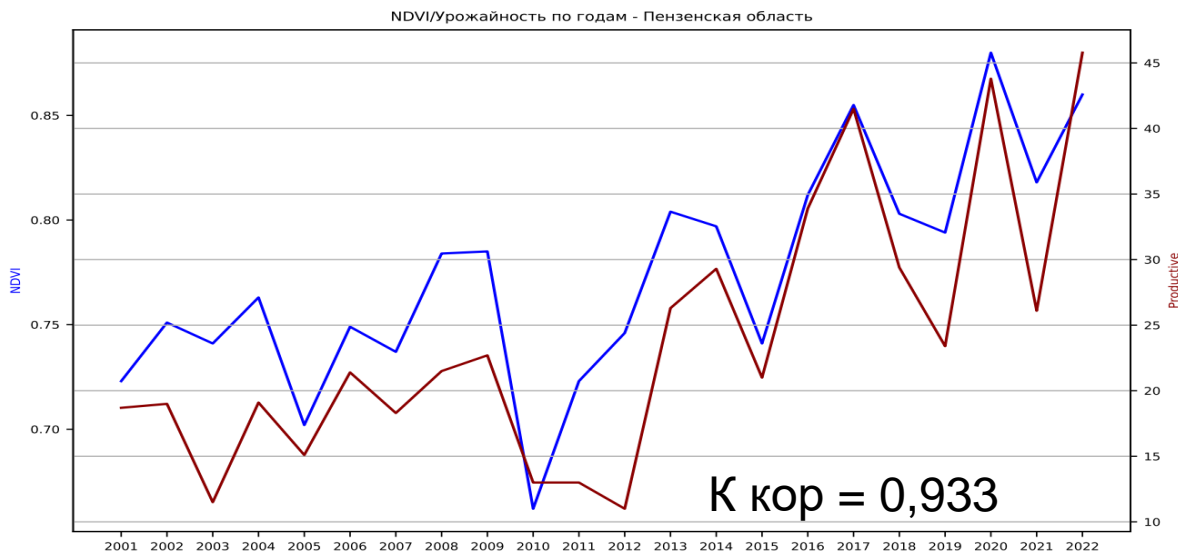
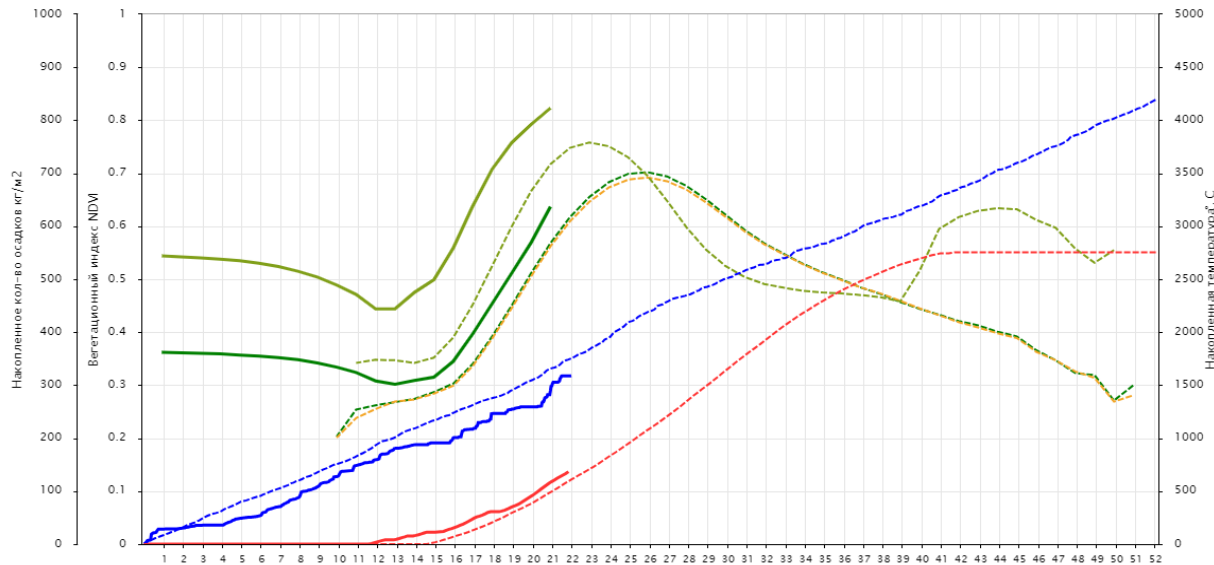
## Исходные данные

- 1) Загрузка графиков NDVI происходит с ресурса BEGA-Science. Для этого программе необходимо сообщить актуальный логин и пароль (<http://sci-vega.ru/>).
- 2) Загрузка данных об урожайности по региону происходит с ресурса ЕМИСС (<https://www.fedstat.ru/indicator/31533>).
- 3) Программа может загружать информации из XLSX файлов, например из бюллетеней о состоянии сельского хозяйства, публикуемых Федеральной Службой Государственной Статистики. (<https://rosstat.gov.ru/folder/11110/document/13277>). Такая возможность необходима, так как на ресурс ЕМИСС могут быть не загружены актуальные данные об урожайности за прошлый год.
- 4) Кроме прочего, программа считывает данные о дополнительных точках в выборку, получаемые с эталонных полей.

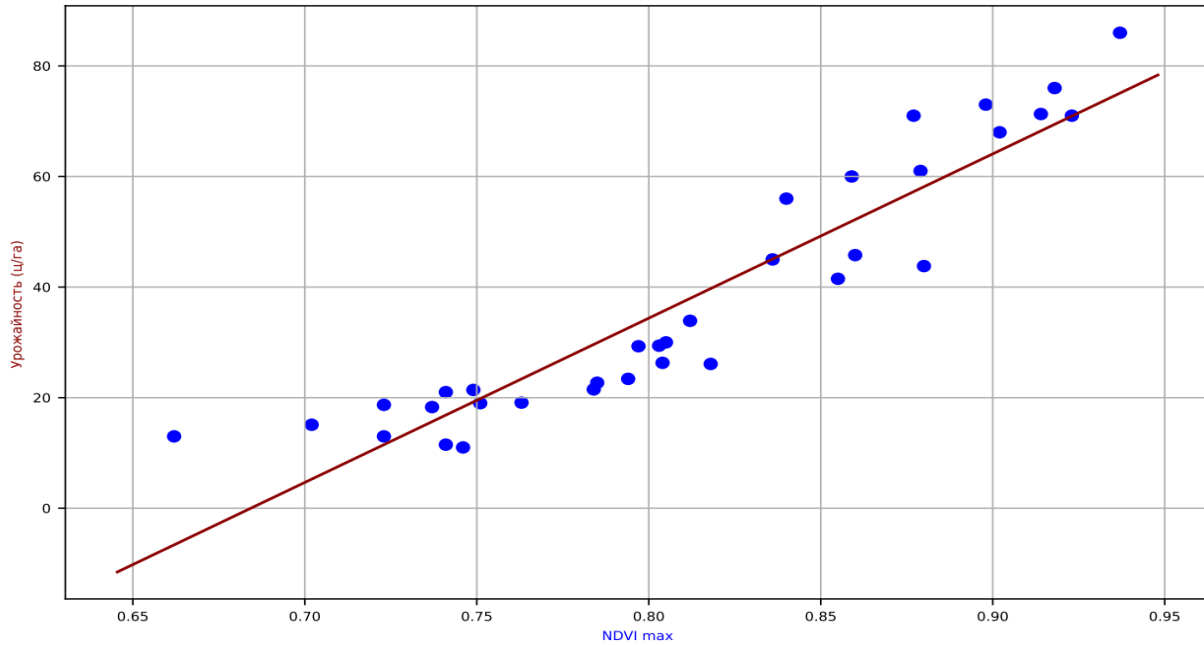


ФЕДЕРАЛЬНАЯ СЛУЖБА ГОСУДАРСТВЕННОЙ СТАТИСТИКИ (РОССТАТ) ГЛАВНЫЙ МЕЖРЕГИОНАЛЬНЫЙ ЦЕНТР	
<b>ВАЛОВЫЕ СБОРЫ И УРОЖАЙНОСТЬ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР ПО РОССИЙСКОЙ ФЕДЕРАЦИИ в 2022 году</b>	
Часть 1	
Москва 2023	
Обложка	Содержание
5	6
17	20
21	22
23	24
25	104(1000001)

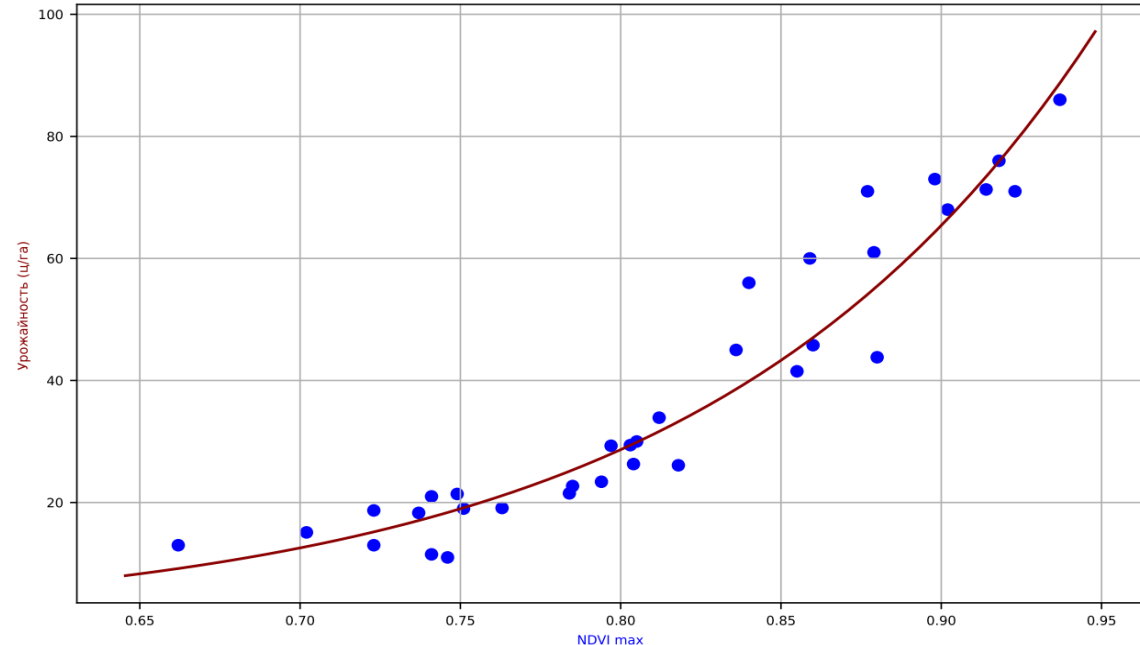
# Процесс обработки статистической информации для Пензенской области. Озимая пшеница. Лето 2023 г.

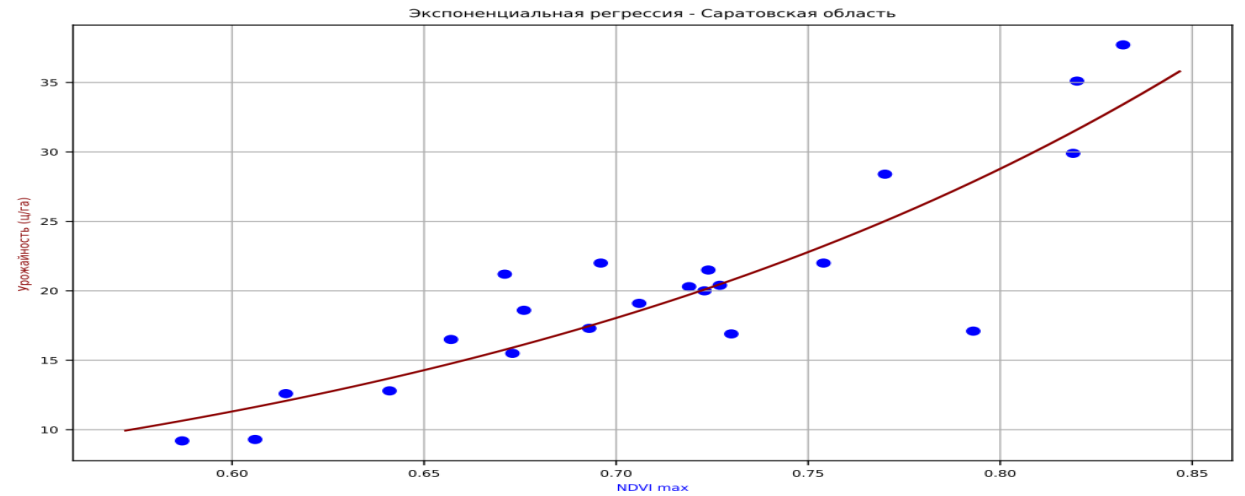
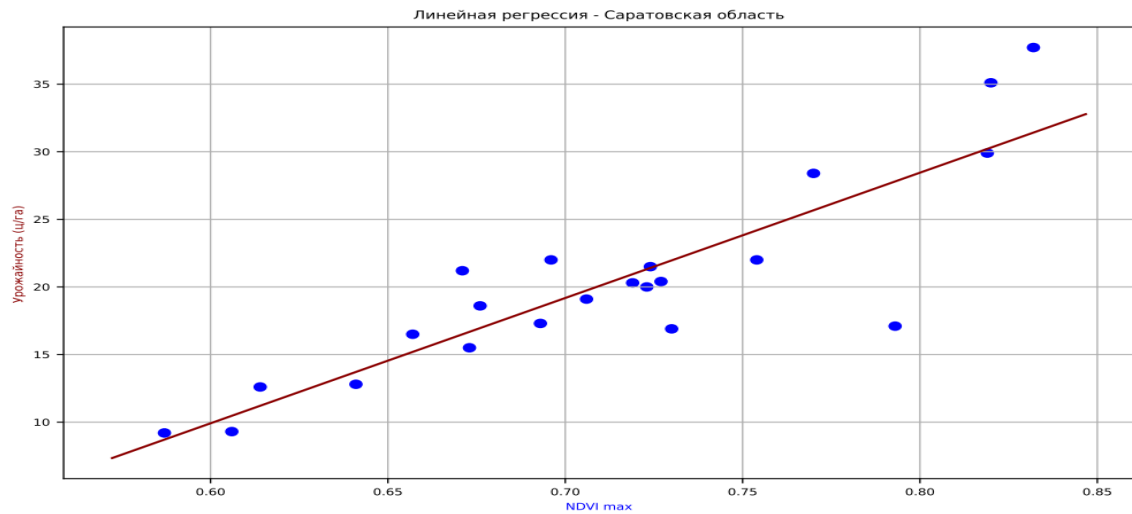
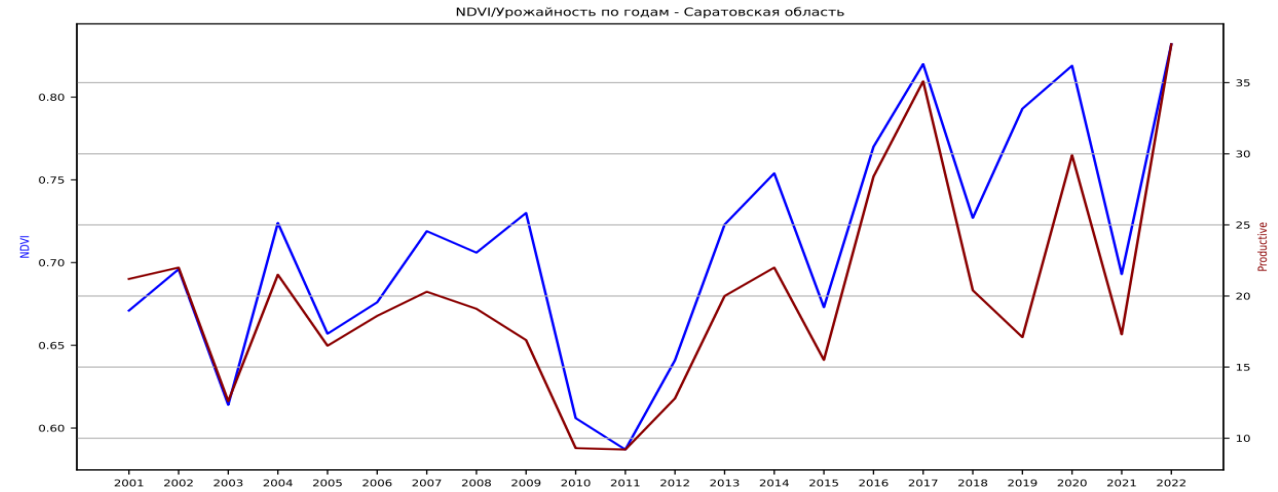
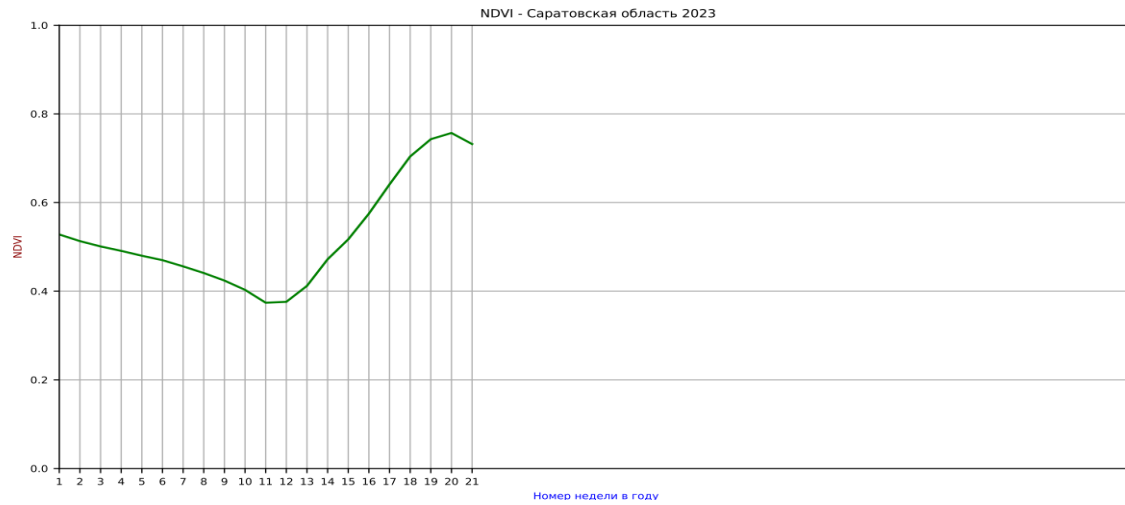


Линейная регрессия - Пензенская область



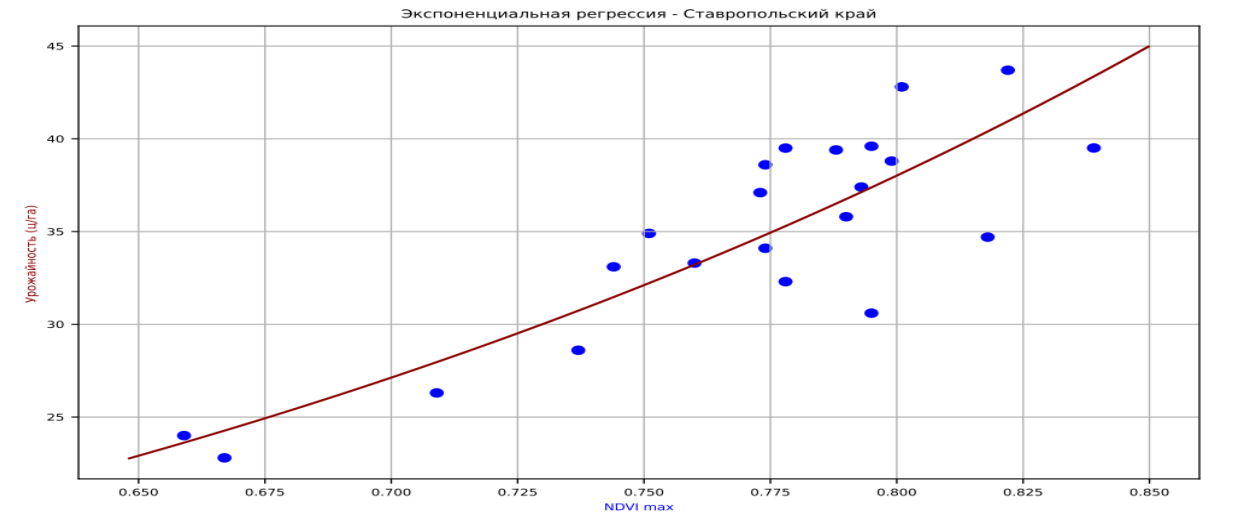
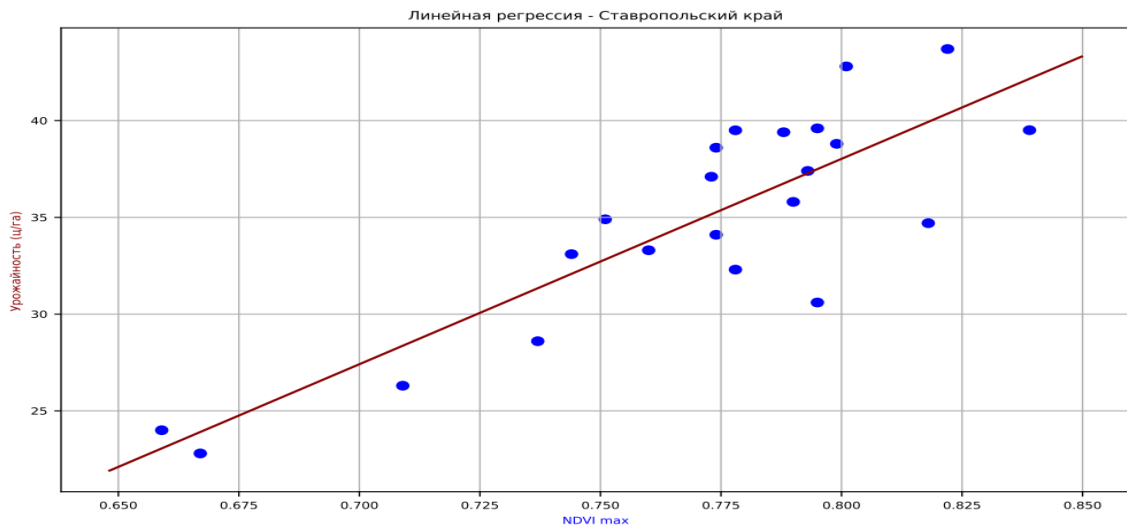
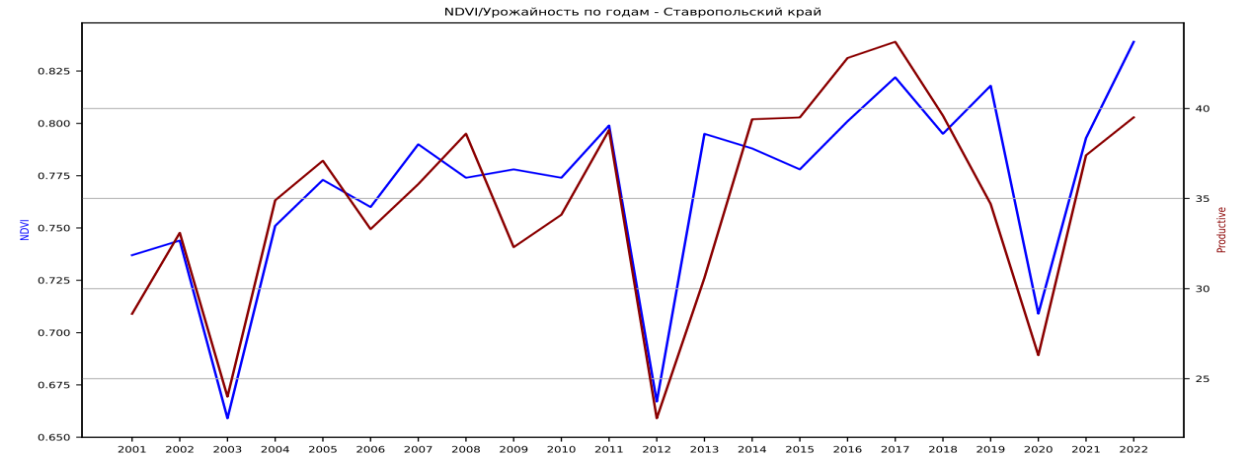
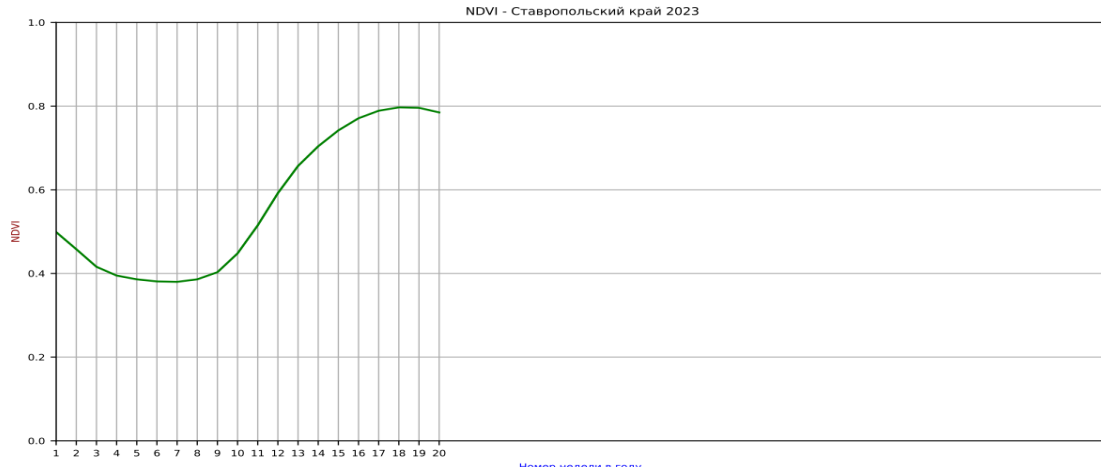
Экспоненциальная регрессия - Пензенская область





Корреляция	Модель аппроксимации	R <sup>2</sup>	MAE	Макс. ошибка	Прогноз (ц/га)
0,87	$-45.703587064870675 + (92.69317127685962) * x$	0,76	2,42	10,70	24,46
	$0.6873333979582398 * (106.53864989741915) ** x$	0,78	2,31	10,76	23,55

# Процесс обработки статистической информации для Ставропольского края. Озимая пшеница. Лето 2023 г.



Корреляция	Модель аппроксимации	R^2	Средняя ошибка (MAE)	Макс. ошибка	Прогноз (ц/га)
0,84	$-46.878734128144686 + (106.12831316300404) * x$	0,72	2,44	6,89	37,71
	$2.5545245656074957 * (29.22869812152441) ** x$	0,76	2,51	6,78	37,63

Данные Productive: Ульяновская область, 2003 найдены в БД программы!

Данные NDVI: Ульяновская область, 2002 найдены в БД программы!

Данные Productive: Ульяновская область, 2002 найдены в БД программы!

Данные NDVI: Ульяновская область, 2001 найдены в БД программы!

Данные Productive: Ульяновская область, 2001 найдены в БД программы!

База данных обновлена!

Введите команду: Обновить БД! / Запустить расчет! / Сформировать отчет! / Показать DATASET! / Завершить работу!

Max NDVI: 0.778

Power model:  $-38.691113356449975 + (80.33803803821522) * x$

Power line:  $1.243136339832733 * (42.17657648672595) ** x$

Sample for calculations: 22

Estimated Regression Productivity (Line model): 23.811880237281464

Estimated Regression Productivity (Power model): 22.84674070341146

Расчеты завершены!

Введите команду: Обновить БД! / Запустить расчет! / Сформировать отчет! / Показать DATASET! / Завершить работу!

Регион: Краснодарский край

Выборка для регрессии: 22

Коэффициент корреляции: 0.7816662388175916

Максимальный NDVI в 2023 году: 0.849

Линейная функция регрессии:  $-110.32805488734962 + (197.38146130827639) * x$

(R<sup>2</sup>): 0.6110021089072395 Средняя ошибка (MAE): 3.856818086881205 Макс. ошибка: 13.243568685993587

Степенная функция регрессии:  $1.7986357404519553 * (58.89452041240386) ** x$

(R<sup>2</sup>): 0.6435664191424573 Средняя ошибка (MAE): 3.9343417052345355 Макс. ошибка: 13.313438523923459

Ожидаемая урожайность по линейной модели в 2023 году: 57.24880576337702

Ожидаемая урожайность по степенной модели в 2023 году: 57.24489989420057

Средняя урожайность за прошлые годы: 51.3811932317583

# Автоматический отчет в формате excel – общая структура

№7

Регион	Показатель	2001	2002	...	2021	2022	2023	Корреляция	Модель аппроксимации	R^2	MAE	Макс. ошибка	Прогноз (ц/га)
Белгородская область	NDVI	0,798	0,763	...	0,87	0,854	0,761	0,857	$-145.0613491049484 + (227.10101442910303) * x$	0,735	4,57	9,87	27,76
	Урожайность	34,6	32,6	...	45	56,24753	-		$0.20409334668290185 * (622.3013756407464) ** x$	0,750	4,38	11,26	27,29
Воронежская область	NDVI	0,764	0,725	...	0,848	0,842	0,792	0,815	$-79.09414669052047 + (140.64302563020914) * x$	0,665	3,88	10,67	32,29
	Урожайность	30,5	29,2	...	29,5	47,63325	-		$0.8177139548796768 * (99.22946572305635) ** x$	0,654	3,69	10,84	31,18
Курская область	NDVI	0,765	0,773	...	0,867	0,867	0,745	0,885	$-145.04602777299533 + (224.476583322168) * x$	0,783	4,47	11,10	22,18
	Урожайность	25,5	28,6	...	43,8	60,5936	-		$0.19739177805843264 * (596.5336797701276) ** x$	0,836	3,73	10,26	23,07
Липецкая область	NDVI	0,766	0,76	...	0,857	0,882	0,738	0,836	$-98.04031599055814 + (165.34871060080812) * x$	0,698	4,31	11,42	23,98
	Урожайность	29,5	31,5	...	37,7	54,59824	-		$0.8123190033004143 * (103.10813791285979) ** x$	0,686	4,09	10,55	24,86

Пензенская область	NDVI	0,898	0,918	0,879	0,902	0,859	0,84	0,805	0,923	0,877	0,836	0,914	0,937
	Урожайность	73	76	61	68	60	56	30	71	71	45	71,3	86



# Автоматический отчет в формате XLSX – пример документа.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	
1	Урожайность сельскохозяйственных культур (в расчете на убранный площадь) (значения показателя за год), центнеров с гектара, Хозяйства всех категорий																																		
2																																			
3	Регион	Показате	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023		Корреляция	Модель а	СКО (R <sup>2</sup> )	Средняя с	Макс. ош	Прогнос			
4	Белгородская область	NDVI	0,798	0,763	0,705	0,799	0,787	0,767	0,801	0,845	0,799	0,766	0,805	0,776	0,786	0,815	0,795	0,832	0,833	0,824	0,865	0,867	0,87	0,854	0,761		0,8575	-145,061	0,73531	4,57985	9,87402	27,7625			
5		Урожайн	34,6	32,6	19,3	27,5	31,8	25,6	31	45,2	31,5	20,7	34,1	31,8	40,7	49,9	37,9	45,1	52,4	44,6	48,6	54	45	56,2475			0,204093	0,75034	4,38059	11,2687	27,2938				
6	Воронежская область	NDVI	0,764	0,725	0,707	0,788	0,734	0,718	0,754	0,814	0,776	0,728	0,752	0,719	0,781	0,807	0,784	0,833	0,834	0,827	0,849	0,854	0,848	0,842	0,792		0,81572	-79,09414	0,6654	3,88719	10,6711	32,2951			
7		Урожайн	30,5	29,2	22,7	24	25,2	20,1	24,9	38,4	28,6	15,4	25,5	26,1	28,9	38,8	29,1	37,4	45,6	35,5	36,6	44,7	29,5	47,6332			0,817713	0,65427	3,69414	10,8416	31,1846				
8	Курская область	NDVI	0,765	0,773	0,695	0,803	0,773	0,725	0,786	0,82	0,808	0,766	0,796	0,801	0,833	0,828	0,788	0,838	0,856	0,826	0,874	0,868	0,867	0,867	0,745		0,88512	-145,046	0,78344	4,47105	11,1087	22,189			
9		Урожайн	25,5	28,6	18,4	24,1	28,1	25	27,3	36,6	34,1	21,7	28,6	30,8	37,7	46,1	31,6	41,4	53,3	45,1	49,5	56,9	43,8	60,5936			0,197391	0,83686	3,73535	10,2662	23,0769				
10	Липецкая область	NDVI	0,766	0,76	0,726	0,801	0,782	0,733	0,799	0,827	0,837	0,786	0,774	0,73	0,846	0,826	0,785	0,853	0,828	0,824	0,863	0,881	0,857	0,882	0,738		0,83605	-98,0403	0,69897	4,31106	11,4238	23,987			
11		Урожайн	29,5	31,5	25,4	28,6	35,4	28,7	33,5	41,8	37,5	20,5	26,2	23,3	36,5	39,1	25,2	40,4	44,9	42,9	42,1	55,4	37,7	54,5982			0,812319	0,68623	4,09184	10,5581	24,862				
12	Орловская область	NDVI	0,781	0,796	0,739	0,821	0,792	0,766	0,801	0,818	0,831	0,785	0,782	0,798	0,841	0,844	0,797	0,852	0,862	0,83	0,875	0,875	0,858	0,876	0,752		0,90386	-130,686	0,81696	2,77857	9,27238	21,3246			
13		Урожайн	27,6	32,9	22,7	26	27,6	28,1	31,3	34,2	35,1	25	23,6	30	38,3	46,5	29,6	37,6	43,9	41,7	43	48,8	42,3	51,7024			0,289538	0,82748	2,48663	8,27356	22,9464				
14	Рязанская область	NDVI	0,779	0,774	0,743	0,829	0,784	0,781	0,779	0,82	0,831	0,758	0,735	0,75	0,835	0,797	0,805	0,845	0,848	0,802	0,823	0,869	0,835	0,87	0,63		0,87784	-117,144	0,7706	3,12479	10,7693	-2,2294			
15		Урожайн	23,3	27,7	21,5	23,3	22,9	28,3	28,9	33,5	32,9	15,9	17,6	18,9	30,3	30,8	29	34	41,7	31,8	32,4	45,1	32,8	47,3306			0,179462	0,76418	2,98996	9,96492	9,96429				
16	Тамбовская область	NDVI	0,766	0,755	0,752	0,788	0,713	0,744	0,756	0,825	0,817	0,738	0,765	0,732	0,845	0,828	0,8	0,839	0,847	0,815	0,848	0,885	0,854	0,874	0,786		0,87861	-90,2513	0,77195	3,27195	7,46998	28,3125			
17		Урожайн	28,1	26,3	23,2	21,5	21,8	25,1	24,4	34,3	29,2	15,7	22,9	20,8	34,8	35,9	29,1	36,7	44,9	35,3	31,4	48,8	31,1	45,9224			0,547993	0,77461	3,1049	7,92003	27,3042				
18	Тульская область	NDVI	0,785	0,794	0,76	0,832	0,796	0,759	0,79	0,835	0,857	0,794	0,763	0,784	0,844	0,823	0,805	0,845	0,86	0,829	0,858	0,872	0,859	0,878	0,69		0,84371	-116,177	0,71185	3,24769	9,88998	7,36532			
19		Урожайн	23,3	27,9	20,1	22,9	23,4	23,4	31,2	31	32,6	19,5	20	23,9	32,3	36	28,4	34,3	37,2	39	33,9	43,4	38,8	48,3884			0,231587	0,71852	3,12189	8,93643	13,7401				
20	Краснодарский край	NDVI	0,851	0,778	0,735	0,785	0,782	0,799	0,827	0,831	0,806	0,823	0,848	0,773	0,847	0,836	0,835	0,838	0,846	0,818	0,843	0,801	0,862	0,86	0,849		0,78167	-110,328	0,611	3,85682	13,2436	57,2488			
21		Урожайн	44,4	47,5	33,8	43,4	46,8	42,9	45,4	55,4	46,1	50,2	55,3	40,1	50,3	55	57,6	58,6	62,1	61,6	59,7	47,8	60	66,3863			1,798635	0,64357	3,93434	13,3134	57,2449				
22	Волгоградская область	NDVI	0,725	0,725	0,553	0,716	0,686	0,662	0,673	0,705	0,722	0,597	0,612	0,67	0,713	0,753	0,721	0,743	0,813	0,751	0,806	0,783	0,769	0,815	0,78		0,80407	-21,2856	0,64652	2,78438	5,42803	28,276			
23		Урожайн	27,5	24,3	16,1	28,5	23,7	22,5	19,8	28,6	21,1	13,4	18,8	18,7	20,4	24,8	21	30,2	32,9	23,3	24,5	28,7	25,9	35,4283			3,149525	0,67961	2,66612	5,92786	28,2811				
24	Ростовская область	NDVI	0,823	0,735	0,612	0,742	0,738	0,752	0,71	0,779	0,764	0,748	0,766	0,719	0,787	0,812	0,785	0,806	0,84	0,802	0,84	0,802	0,838	0,854	0,842		0,7974	-43,0985	0,63585	3,30437	9,31733	38,5507			
25		Урожайн	30,4	30,7	19,6	35,9	30,4	29,3	22,2	34,1	26,8	27,3	29,2	23,8	23,9	33,4	32,5	39	44,5	36	35,6	36,5	39,6	44,8689			2,616771	0,67029	3,12239	8,68865	38,8698				
26	Ставропольский край	NDVI	0,737	0,744	0,659	0,751	0,773	0,76	0,79	0,774	0,778	0,774	0,799	0,667	0,795	0,788	0,778	0,801	0,822	0,795	0,818	0,709	0,793	0,839	0,797		0,84765	-46,8787	0,71851	2,44477	6,89327	37,7055			
27		Урожайн	28,6	33,1	24	34,9	37,1	33,3	35,8	38,6	32,3	34,1	38,8	22,8	30,6	39,4	39,5	42,8	43,7	39,6	34,7	26,3	37,4	39,506			2,554524	0,7568	2,506	6,77908	37,6323				
28	Республика Татарстан	NDVI	0,757	0,792	0,783	0,769	0,727	0,787	0,772	0,797	0,752	0,618	0,805	0,721	0,78	0,76	0,767	0,784	0,821	0,778	0,826	0,831	0,707	0,823	0,722		0,81476	-76,9970	0,66383	3,88191	8,8435	22,3409			
29		Урожайн	36	37,9	29,5	25,3	26,2	32,2	32,3	35,4	33,1	10,3	30,9	17,7	26,3	22	20,1	31,6	36,7	28,9	29,4	42,2	16,1	39,0319			0,249869	0,75	3,63183	10,5062	20,5855				
30	Пензенская область	NDVI	0,723	0,751	0,741	0,763	0,702	0,749	0,737	0,784	0,785	0,662	0,723	0,746	0,804	0,797	0,741	0,812	0,855	0,803	0,794	0,88	0,818	0,86	0,774		0,93335	-203,322	0,87114	6,78008	19,6245	26,6537			
31		Урожайн	18,7	19	11,5	19,1	15,1	21,4	18,3	21,5	22,7	13	13	11	26,3	29,3	21	33,9	41,5	29,4	23,4	43,8	26,1	45,7832			0,038987	0,90246	4,63563	16,9072	23,126				
32	Самарская область	NDVI	0,689	0,767	0,708	0,728	0,671	0,707	0,737	0,729	0,763	0,609	0,688	0,644	0,755	0,788	0,729	0,779	0,845	0,766	0,79	0,842	0,722	0,87	0,759		0,8948	-67,5524	0,80067	3,01555	8,26915	23,5841			
33		Урожайн	20,5	21,3	12	15,2	12,9	14,2	19,7	19,1	20,4	12,1	17,7	11,7	22	25,9	18,3	26,9	35,7	25,3	20,5	37,1	20,5	45,1815			0,418491	0,83246	2,28637	5,86355	22,0229				
34	Саратовская область	NDVI	0,671	0,696	0,614	0,724	0,657	0,676	0,719	0,706	0,73	0,606	0,587	0,641	0,723	0,754	0,673	0,77	0,82	0,727	0,793	0,819	0,693	0,832	0,757		0,87459	-45,7035	0,76491	2,42138	10,7021	24,4651			
35		Урожайн	21,2	22	12,6	21,5	16,5	18,6	20,3	19,1	16,9	9,3	9,2	12,8	20	22	15,5	28,4	35,1	20,4	17,1	29,9	17,3	37,708			0,687333	0,78945	2,31198	10,76</					

## Яровые зерновые и зернобобовые культуры

### Алтайский край

Выборка для регрессии: 22. Данные с 2023 по 2001 год.

Коэффициент корреляции урожай (Яровые зерновые и зернобобовые культуры) и максимального значения вегетационного индекса NDVI: 0.584

Максимальное значение NDVI, измеренное в 2023 году: 0.652



Рис. Сопоставление урожайности и максимальных значений вегетационного индекса.



Рис. График вегетационного индекса NDVI измеренный в 2023 году.

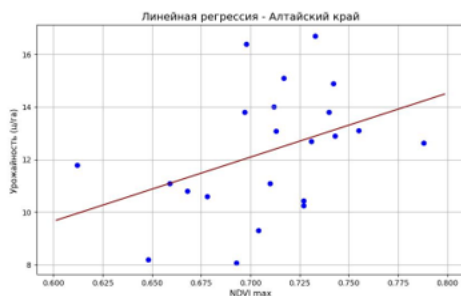


Рис. Линейная модель аппроксимации зависимости средней урожайности в регионе от max\_NDVI.

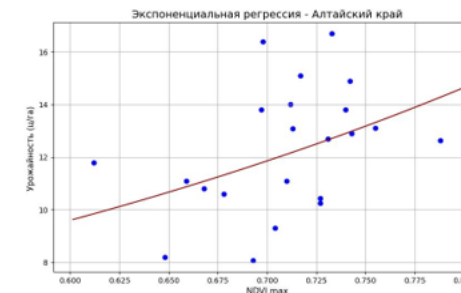


Рис. Экспоненциальная модель аппроксимации зависимости средней урожайности в регионе от max\_NDVI.

Таблица – Параметры аппроксимации.

	f(x)	R <sup>2</sup>	Средняя ошибка (ц/га)	Максимальная ошибка (ц/га)	Ожидаемая урожайность в 2023 году (ц/га)
Линейная модель	$4.93+(24.31)*x$	0.159	1.744	4.355	10.927
Экспоненциальная модель	$2.7*(8.26)^x$	0.17	1.739	4.592	10.715

Средняя урожайность за прошлые годы: 12.31

Пример DOCX файла с отчетом о произведенных расчетах, созданный программой автоматически. Сформированный документ сохраняется в папку “Word report”. Яровые зерновые и зернобобовые. Лето 2023 г.

1. Изучены существующие методы оценки состояния посевов.
2. Разработаны 2 программы для автоматизированного сбора статистической и спутниковой информации для ресурсов Вега (<http://sci-vega.ru/>) и ЕМИСС (<https://www.fedstat.ru/indicator/31533>).
3. Разработана программа для автоматизированной обработки информации и вычисления ожидаемой урожайности.
4. На данный момент программа имеет консольный интерфейс. Начальные указания к работе необходимо прописывать напротив соответствующих ключей в отдельном .json файле.
5. Программа автоматически формирует отчет в формате excel. Также программа создает набор графиков в векторном формате pdf для каждого региона.
6. Программа может быть использована для поиска культур, урожайность которых коррелирует с максимальными значениями NDVI по разным маскам.

### Дальнейший план:

1. Написать программы-парсеры для нескольких других сайтов статистической информации с целью увеличить выборки для построения моделей.
2. Разработать полноценный оконный user-friendly интерфейс.
3. Попытаться использовать другие методы обработки данных, учитывающих не только NDVI, но и другие продукты ресурса Вега.