

Порайонные прогнозы урожайности озимой пшеницы для двух областей России: сравнение точности моделей машинного обучения и классических подходов с использованием алгоритма локальной адаптации

Авторы (ИКИ РАН):
Мещеряков А.В., Кашницкий А.В., Трошко К.А., Денисов П.В., Толпин В.А.

Двадцать четвертая международная конференция "СОВРЕМЕННЫЕ ПРОБЛЕМЫ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ ИЗ КОСМОСА"

Пшеница является одной из основных зерновых культур в мире и обеспечивает условия продовольственной безопасности. Прогнозы урожайности озимой пшеницы имеют большое прикладное значение для регионов России, где озимые зерновые культуры являются доминирующими культурами: в регионах Центральной России, а также на юге России, - в областях с мягкими зимами и достаточным снежным покровом, защищающим всходы от вымерзания. В целом, величина урожайности культуры представляет собой интегральный показатель состояния сельскохозяйственных земель.

В данной работе мы исследовали ряд современных подходов к ретроспективному определению урожайности озимой пшеницы на районном уровне на примере двух областей России (Пензенская и Ростовская области), которые находятся в разных природно-климатических условиях. Нами были использованы полученные на основе прибора MODIS спутниковые данные за 2013-2025 гг. (Толпин и др., 2014), а также данные реанализа погоды, собранные в системе Вега-Science (Loupian et al., 2022), и доступная порайонная статистика урожайности Росстата.

Модели урожайности **озимой пшеницы** (по районам) в ИКИ РАН по данным Vega-Science

Временные особенности используемых данных

1. **Базовая модель по NDVImax**
робастная линейная регрессия yield(NDVImax) с учетом районных и годовых трендов, прогноз по данным текущего года

агрегация за сезон вегетации - NDVImax

2. Neural Rabbit (Сычков и Трошко 2024)
нейросеть с на ежедневных интерполированных рядах NDVI+метео, прогноз по данным текущего года и средних многолетних рядов

ежедневные интерполированные ряды NDVI и реанализа погоды

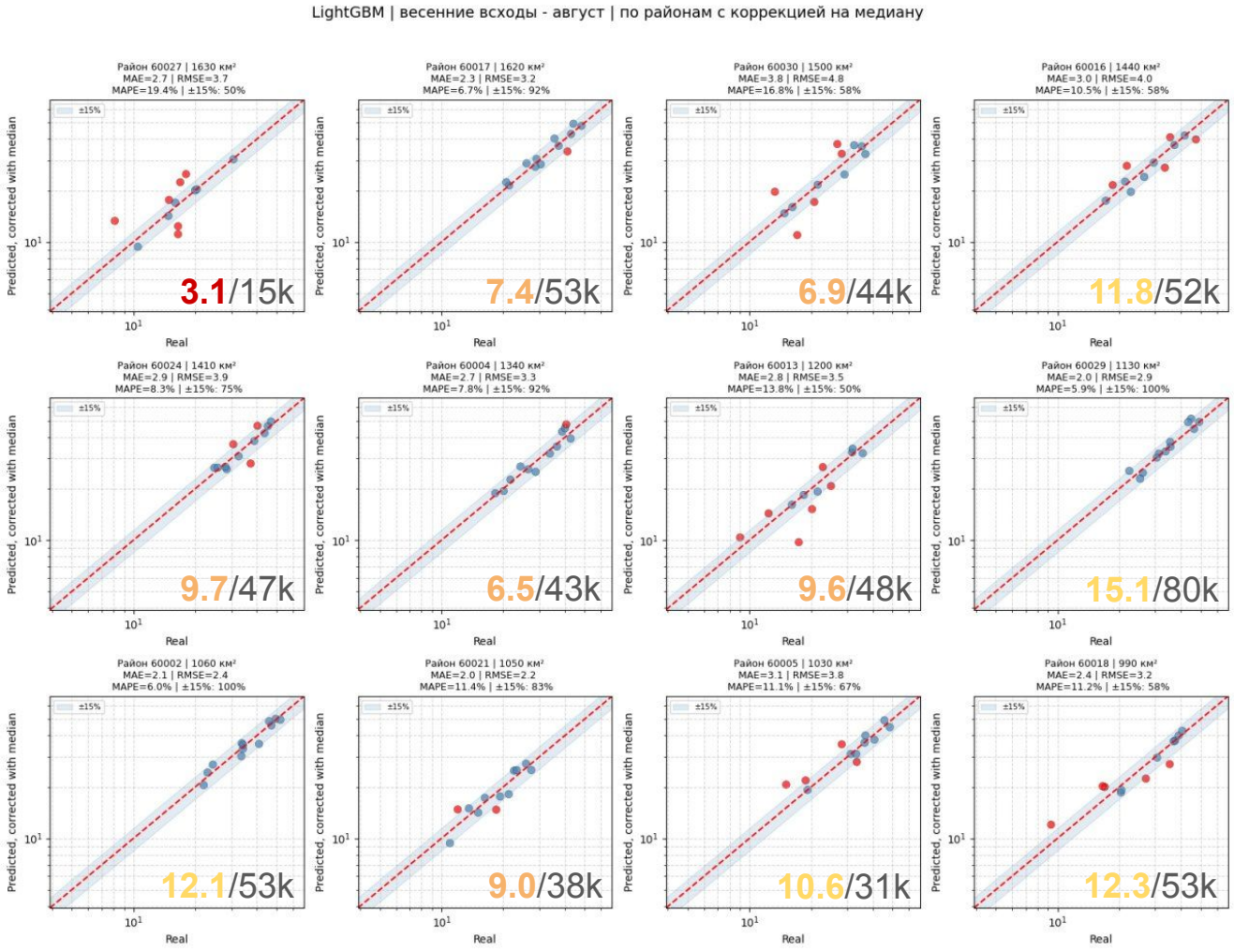
3. Подекадная ML-модель
градиентный бустинг на подекадных рядах NDVI+признаки из метео (см. признаки в работе Клещенко 2025), прогноз по данным текущего года с калибровкой по прошлому году

подекадные интерполированные ряды NDVI и реанализа погоды

Все три модели демонстрируют высокую точность для благоприятных лет развития прсевов. При наступлении неблагоприятных погодных условий для развития культуры в сезоне вегетации, информация о погоде (соотнесенная с наступлением фенофаз озимой пшеницы) в модели (2) обеспечивает ей более высокую точность прогноза урожайности, чем дает классическая модель (1), даже с учетом ее локальной адаптации. В модели (3) наряду с временными рядами NDVI и погодных параметров на год прогноза, используются также усредненные временные ряды за несколько прошлых лет. Это создает модели (3) сложности при резких колебаний погодных условий на масштабе 1 года. Модель (2) демонстрирует высокую точность прогноза урожайности как в благоприятные так и неблагоприятные годы с точки зрения развития посевов озимой пшеницы и в разных погодных условиях за сезон вегетации.

Далее будут приведены некоторые результаты подекадной ML-модели (2)

Точность прогнозов для районов Пензенской области на 2013-2024



Площадь посевов озимой пшеницы (2025):

>20k

10-20k

5-10k

<5k

Проблема систематических ошибок прогнозов порайонной урожайности: сравнение Пензенской и Ростовской областей на 2025

region	2023_gt	2024_gt	2025_gt	S2025	id_distr	d2023(cros	d2024(cros	d2025(cros
Пензенская область	40,47	34,7	43,45	18118	**60001**	12.89	-2.14	-5
Пензенская область	45,93	34,23	47,74	12066,5	**60002**	15.11	0.51	-3.79
Пензенская область	43,27	36,24	43,62	31870	**60003**	17.41	27.65	4.74
Пензенская область	40,66	36,6	41,47	6522,2	**60004**	21.46	-11.06	0.53
Пензенская область	36,49	40,69	41,71	10610	**60005**	-7.19	-7.68	-1.85
Пензенская область	27,24	26,64	33,77	5249	**60008**	25.94	-17.47	-6.47
Пензенская область	26,34	28,35	39,28	18498	**60010**	41.54	18.67	-7.05
Пензенская область	48,9	34,37	44,6	8834,7	**60011**	-13.03	-11.28	-8.3
Пензенская область	45,09	34,05	39,95	21127,4	**60012**	13.15	-1.43	-1.68
Пензенская область	31,87	25	32,49	9618	**60013**	31.91	3.78	13.39
Пензенская область	42,92	32,26	44,13	21932,5	**60014**	10.86	8.4	-5.24
Пензенская область	30,76	33,29	41,71	5957	**60015**	29.84	-14.9	-22.86
Пензенская область	35,46	33,43	41,2	11818,4	**60016**	23.44	-15.13	1.2
Пензенская область	44,04	41,18	45,77	7394,9	**60017**	4.09	-14.23	-13.22
Пензенская область	36,91	30,62	38,96	12315,4	**60018**	20.11	7.93	2.8
Пензенская область	40,89	34,77	36,67	21869,7	**60019**	12.63	1.55	9.98
Пензенская область	40,54	36,96	35,68	6937	**60020**	5.97	-10.21	-3.3
Пензенская область	27,42	19,33	27,46	9019	**60021**	6.41	17.66	27.96
Пензенская область	42,43	40,57	34,96	15208	**60022**	5.99	-5.75	10.82
Пензенская область	41,14	21,01	35,2	2097	**60023**	-15.75	18.94	2.02
Пензенская область	40,3	38,88	38,15	9671	**60024**	22.11	-3.05	23.12
Пензенская область	45,04	35,53	40,5	28582	**60025**	3.77	-6.03	11.84
Пензенская область	41,13	31,85	39,21	14359	**60026**	15.46	11.24	-3.68
Пензенская область	30,76	20,37	29,92	3096	**60027**	9.17	20.79	3.89
Пензенская область	42,13	40,49	43,39	8735	**60028**	-6.32	-20.5	-17.46
Пензенская область	43,45	35,51	46,41	15066,3	**60029**	18.71	7.66	5.38
Пензенская область	35,37	29,11	39,07	6866,2	**60030**	8.94	-17.97	-12.94

Зелеными ячейками в последнем столбце показаны прогнозы урожайности с точностью <10% на данных 2025 года для Пензенской и Ростовской областей

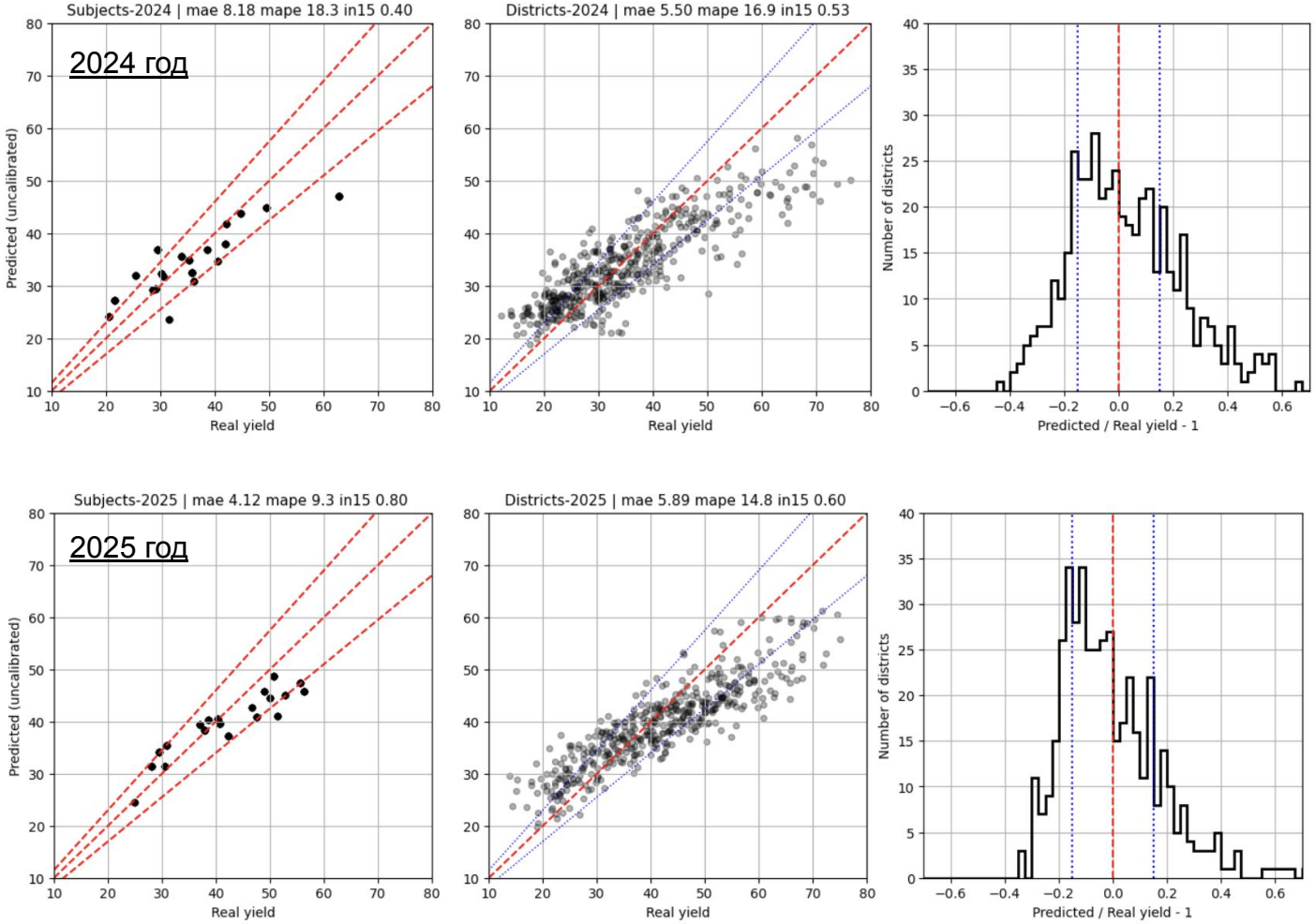
region	2023_gt	2024_gt	2025_gt	S2025	id_distr	d2023(cros	d2024(cros	d2025(cros
Ростовская область	50	35	30,2	20365,19	62002	4.56	3.14	17.09
Ростовская область	48,49	41,1	25,5	25681,08	62003	7.95	1.21	4.09
Ростовская область	42,91	23,8	21,3	58771,66	62004	4.81	19.93	55.3
Ростовская область	41,78	21,6	38,1	33808,76	62005	-5.74	8.81	-9.18
Ростовская область	40,56	25,8	33,5	35145,71	62006	-4.42	-16.99	3.7
Ростовская область	47,4	48,1	26,8	40363,07	62007	8.13	0.35	7.83
Ростовская область	42,84	29,6	21,6	19811,93	62008	6.38	4	30.47
Ростовская область	41,44	31,5	34,4	110822,7	62021	0.96	-9.5	7.73
Ростовская область	51,29	51,5	23,2	60880,6	62022	2.87	1.76	63.51
Ростовская область	36,26	26,6	36,9	89547,55	62023	-3.36	0.53	-8.35
Ростовская область	48,38	53,6	24,8	114149,3	62024	10.78	9.48	76.65
Ростовская область	41,31	34,2	35,2	189628,3	62025	1.14	1.17	25.31
Ростовская область	50,15	45,9	25,2	58297,29	62026	5.99	10.83	59.04
Ростовская область	41,75	23,2	23,1	54718,46	62027	1.67	-0.03	22.2
Ростовская область	47,16	22,4	32,3	77403,61	62028	-5.56	7.95	15.1
Ростовская область	48,63	33,1	23,6	69012,02	62029	-6.23	-19.56	10.12
Ростовская область	43,54	19,9	29,2	56060,37	62030	14.19	22.88	28.49
Ростовская область	60,49	29,8	15,2	30708,33	62031	-16.25	-14.32	85.29
Ростовская область	49,48	41,2	24,4	54226,37	62032	0.88	5.46	7.06
Ростовская область	53,77	32,4	13,8	65742,79	62033	-9.68	-8.01	93.92
Ростовская область	43,64	17,6	32,4	81928,38	62034	0.79	22.1	7.43
Ростовская область	38,24	21,5	24,3	55848,56	62035	5.11	3.66	32.28
Ростовская область	43,78	25,8	23,2	74039,25	62036	-0.3	-2.22	28.07
Ростовская область	60,1	45,1	24,2	29165,4	62037	-15.81	-24.56	61.67
Ростовская область	55,96	42,4	21,8	76413,02	62038	-0.36	0.61	71.25
Ростовская область	40,15	22,9	32,3	46004,22	62039	-5.89	-7.11	10.1
Ростовская область	48,96	30,8	22,2	61680,66	62040	2.8	-6.27	51.47
Ростовская область	46,59	41,3	33,3	117204,1	62041	-9.63	-16.8	28
Ростовская область	53,26	59,8	30,2	83692,39	62042	0.27	-4.61	61.29
Ростовская область	46,93	43,6	30,2	76543,9	62043	-5.35	-2.91	2.26
Ростовская область	37,98	30,2	38,3	89373,6	62044	-7.61	-6.69	0.66
Ростовская область	51,47	29,6	20,6	60695,63	62045	-7.31	-3.84	45.09
Ростовская область	47,25	51,3	31,5	146257,1	62046	3.98	3.47	37.6
Ростовская область	46,95	37,2	23	33201,52	62047	12.72	-8.95	3.11
Ростовская область	33,55	20,5	30	24387,95	62048	8.42	9.52	3.4
Ростовская область	42,76	22,3	26,8	69999,72	62049	2.99	3.33	30.95
Ростовская область	42,22	24,7	21,6	66807,72	62050	6.52	-2.15	24.07
Ростовская область	49,87	30,5	19,7	31348,78	62051	-0.3	3.99	61.52
Ростовская область	55,61	58,8	26,6	94501,45	62052	-1.06	-6.89	66.77
Ростовская область	44,55	30,7	27,9	44745,17	62053	-4.46	-13.51	0.52
Ростовская область	45,33	21,6	34,4	59825,68	62054	-3.25	26.9	13.38
Ростовская область	43,39	23	37	33214,68	62055	-8.27	-4.67	0.54

Точность прогнозов для 20 областей и их районов Центральной России

Распределения для прогнозов по областям (слева) и районов (в центре), Y:прогноз X: фактическая урожайность; справа - распределение относительных отклонений прогноза от факта

Здесь мы видим схожее поведение прогнозов - для обоих лет прогнозы идут под ~ одним наклоном к диагонали (+-15% диапазон ошибок показан линиями). Это явление называется “attenuation bias” и встречается в литературе:

Ting, Yuan-Sen. 2025. “**Why Machine Learning Models Systematically Underestimate Extreme Values.**” The Open Journal of Astrophysics 8 (July). https://doi.org/10.33232/001c.142224
https://arxiv.org/pdf/2412.05806

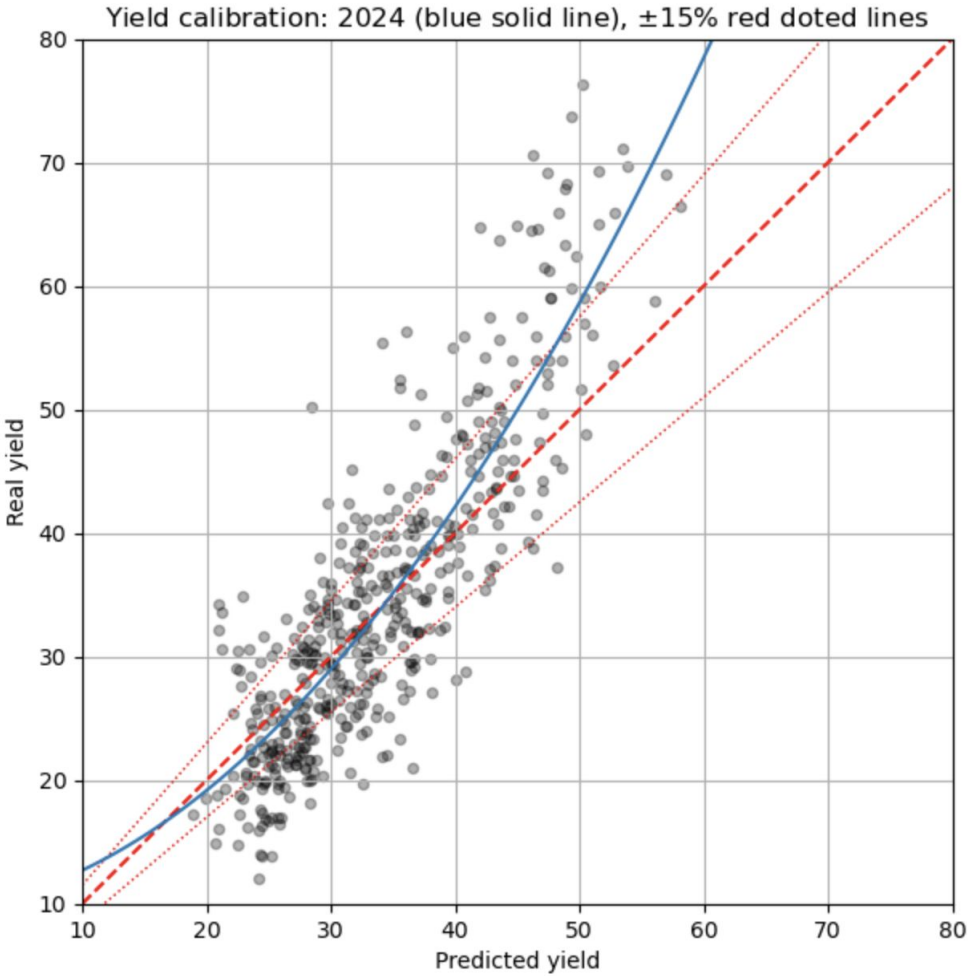


Калибровка прогнозов урожайности по 2024 году

Калибровка районной урожайности выполняется как показано на графике.
X: прогноз урожайности по районам в 2024
Y: фактическая урожайность по районам в 2024

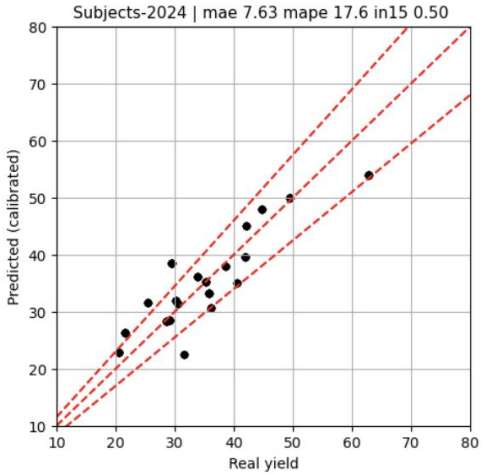
Тоесть для 2024 года мы построили зависимость - фактическая порайонная урожайность от значений урожайности предсказываемых моделью (см. точки на рисунке - районы 20 субьектов с площадью посевов озимой пшеницы >5000). Далее эта зависимость аппроксимируется полиномом 2-й степени (синия сплошная линия на рисунке) и применяется к прогнозам целевого 2025 года.

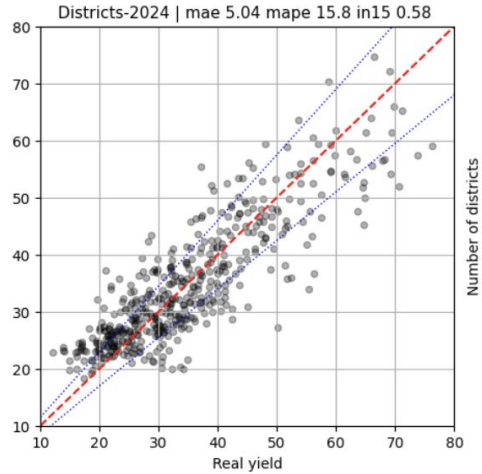
Хочу отметить, что калибровка это временное решение - как только модель справится с систематикой в дополнительной калибровке отпадет необходимость.

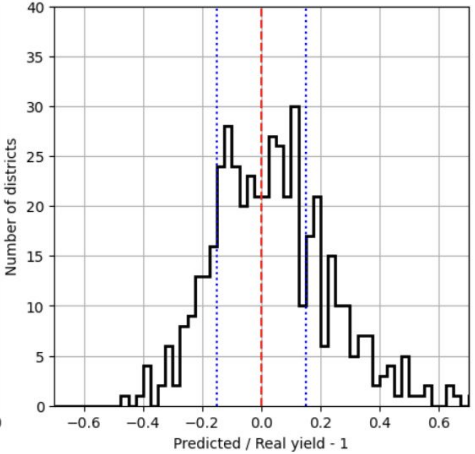


Результат калибровки прогнозов по 2024 году

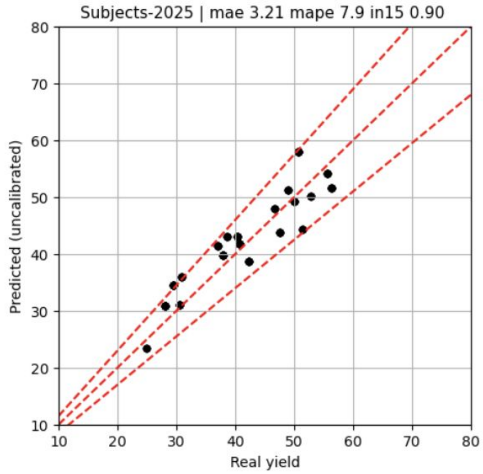
2024 год

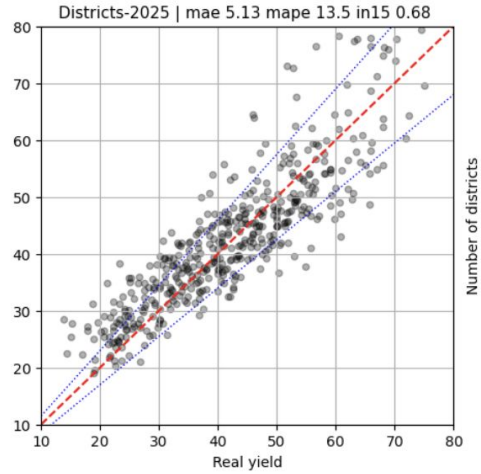


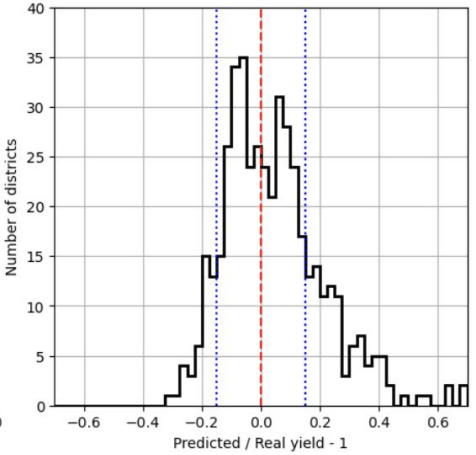




2025 год







После калибровки прогнозы за 2025 год не показывают систематики!

Литература

- Клещенко, А. Д. Оценка ожидаемой средней районной урожайности озимой пшеницы с использованием полносвязной нейронной сети / А. Д. Клещенко, О. В. Савицкая, Я. А. Вдовина // Гидрометеорологические исследования и прогнозы. – 2025. – № 3(397). – С. 132-145. – DOI 10.37162/2618-9631-2025-3-132-145. – EDN GLLJRY
- Лупян Е.А., Прошин А.А., Бурцев М.А., Кашницкий А.В., Балашов И.В., Барталев С.А., Константинова А.М., Кобец Д.А., Мазуров А.А., Марченков В.В., Матвеев А.М., Радченко М.В., Сычугов И.Г., Толпин В.А., Уваров И.А. Опыт эксплуатации и развития центра коллективного пользования системами архивации, обработки и анализа спутниковых данных (ЦКП «ИКИ-Мониторинг») // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 3. С. 151-170. DOI: 10.21046/2070-7401-2019-16-3-151-170
- Loupian, E.; Burtsev, M.; Proshin, A.; Kashnitskii, A.; Balashov, I.; Bartalev, S.; Konstantinova, A.; Kobets, D.; Radchenko, M.; Tolpin, V.; et al. Usage Experience and Capabilities of the VEGA-Science System. // Remote Sens. 2022, 14, 77. https://doi.org/10.3390/rs14010077
- Сычков А.А., Трошко К.А. Разработка нейросетевой модели анализа спутниковых данных для оценки урожайности сельскохозяйственных культур в районах РФ // Материалы 22-й Международной конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва: ИКИ РАН, 2024. С. 222. DOI 10.21046/22DZZconf-2024a
- Толпин В.А., Лупян Е.А., Барталев С.А., Плотников Д.Е., Матвеев А.М. Возможности анализа состояния сельскохозяйственной растительности с использованием спутникового сервиса «ВЕГА» // Оптика атмосферы и океана. 2014. Т. 27, № 7. С. 581–586.
- Loupian E., Burtsev M., Proshin A. et al. Usage experience and capabilities of the VEGA-Science system // Remote Sensing. 2022. Vol. 14. No. 1. Art. 77. DOI: 10.3390/rs14010077

Спасибо за внимание!