

Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»
Уральский энергетический институт
Кафедра Автоматизированных электрических систем

На правах рукописи



Сафаралиев Муродбек Холназарович

**РАЗРАБОТКА МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ
ЭЛЕКТРОПОТРЕБЛЕНИЯ И ГЕНЕРАЦИИ ГЭС НА СРЕДНЕСРОЧНУЮ
ПЕРСПЕКТИВУ В ИЗОЛИРОВАННЫХ ЭНЕРГОСИСТЕМАХ**

Специальность 2.4.3. Электроэнергетика

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель:
доктор технических наук, профессор
Кокин Сергей Евгеньевич

Екатеринбург – 2022

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
ГЛАВА 1 ОЦЕНКА СОСТОЯНИЯ ЭНЕРГЕТИЧЕСКИХ РЕСУРСОВ ГОРНО- БАДАХШАНСКОЙ АВТОНОМНОЙ ОБЛАСТИ	13
1.1 Описание объекта исследования	13
1.2 Мургабские районные электрической сети (РЭС)	14
1.3 Потенциал альтернативных источников энергии ГБАО	17
1.3.1 Гидроэнергетические и водные ресурсы	17
1.3.2 Солнечные ресурсы.....	20
1.3.3 Ветровые ресурсы	21
1.4 Влияние ДЭС в ИЭС ГБАО.....	22
Выводы по главе 1.....	26
ГЛАВА 2 АНАЛИЗ МЕТОДОВ СРЕДНЕСРОЧНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ И ГЕНЕРАЦИИ	27
2.1 Классификация методов прогнозирования электропотребления и генерации	27
2.2 Математические методы, используемые для среднесрочного прогнозирования электропотребления и генерации.....	29
2.2.1 Статистические методы.....	29
2.2.2 Методы прогнозирования, основанные на машинном обучении	32
2.3 Влияние метеорологических факторов и изменения климата на выбор модели прогнозирования	36
2.4 Сравнение методов прогнозирования электропотребления и генерации	38
Выводы по главе 2.....	40

ГЛАВА 3 РАЗРАБОТКА МОДЕЛИ СРЕДНЕСРОЧНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ И ГЕНЕРАЦИИ С УЧЕТОМ МЕТЕОРОЛОГИЧЕСКИХ ФАКТОРОВ	41
3.1 Среднесрочная модель прогнозирования энергопотребления	41
3.1.1 Описание выборки данных	42
3.1.2 Постановка задачи прогнозирования	44
3.1.3 Кластеризация метеорологических условий	45
3.1.4 Прогнозирование электропотребления	46
3.1.5 Нейросетевая модель и результаты	49
3.2 Среднесрочная модель прогнозирования притока воды для планирования генерации	52
3.2.1 Постановка задачи и исходные данные	53
3.2.2 Используемый метод и результаты	55
3.2.3 Описание экспериментов	56
3.2.4 Модели на базе машинного обучения и результаты	58
3.2.5 Анализ полученных результатов	62
3.3 Среднесрочная модель прогнозирования генерации гидроэлектростанциями	63
3.3.1 Постановка задачи и исходные данные	64
3.3.2 Анализ данных и формирование выборки для применения моделей машинного обучения	65
3.3.3 Используемые модели машинного обучения и результаты	67
Выводы по главе 3	73
ГЛАВА 4 КОМПЛЕКСНОЕ ПЛАНИРОВАНИЕ ГРАФИКА ГЕНЕРАЦИИ НА СРЕДНЕСРОЧНОЮ ПЕРСПЕКТИВУ	74
4.1 Планирование генерации	74

4.2 Анализ и предварительная обработка данных	75
4.3 Прогнозирование и критерии оптимизации	80
4.4 Оценка влияния прогноза на планирование генерации ДЭС	84
4.5 Расчет экономических и экологические показатели	90
Выводы по главе 4.....	94
ЗАКЛЮЧЕНИЕ	96
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ	99
СПИСОК ЛИТЕРАТУРЫ.....	100
СПИСОК РИСУНКОВ.....	115
СПИСОК ТАБЛИЦ	117
ПРИЛОЖЕНИЕ АКТ ВНЕДРЕНИЯ РЕЗУЛЬТАТОВ ДИССЕРТАЦИОННОЙ РАБОТЫ.....	119

ВВЕДЕНИЕ

Актуальность темы исследования. С принятием Парижского соглашения 2015 года страны мира поставили цель резко сократить выбросы парниковых газов. Для решения этой амбициозной задачи предложен ряд мер [1,2], ведущую роль среди которых должно сыграть внедрение, распространение и наконец полный переход к выработке электроэнергии с помощью возобновляемых источников энергии (ВИЭ). Так, страны Тихоокеанского, Атлантического и Индийского регионов, для которых повышение уровня мирового океана приведет к затоплению территорий, стремятся к 2050 году полностью заместить традиционные источники энергии ВИЭ. Не являются исключением и континентальные страны, природные условия которых благоприятствуют отказу от тепловой генерации. Так Республика Таджикистан, обладая значительными гидроэнергетическими ресурсами, стремится к 2030 году вырабатывать 100% электроэнергии с помощью ВИЭ.

Следует понимать, что развитие зеленой энергетики не ограничивается задачами уменьшения углеродного следа. Тенденцией последних десяти лет стало применение ВИЭ, как установок распределенной генерации в районах децентрализованного электроснабжения и автономных изолированных энергосистемах. Потребители электроэнергии, проживающие на островах, высокогорьях, а также работающие в труднодоступных районах предприятия добывающей промышленности не могут быть обеспечены электроэнергией от объединенной энергосистемы. Строительство электрических связей с этими областями оказывается экономически нецелесообразным либо из-за сравнительно небольшого внутреннего спроса на электроэнергию, либо, как в последнем случае, временности потребления: когда добываемый ресурс иссякнет, месторождение будет покинуто и построенные линии станут бесполезны. Электроснабжение таких удаленных потребителей обеспечивается собственными малыми системами генерации и распределения электроэнергии, которые вместе образуют изолированную энергосистему (ИЭС). Электроэнергия в ИЭС традиционно вырабатывается с помощью дизельных генераторов, поскольку они обеспечивают

маневренную генерацию и при этом мобильны в установке [3]. Трудности связаны с дизельным топливом. Кроме ущерба местной экологии, проявляется транспортная проблема: из-за удаленности районов доставка топлива кратно увеличивает его стоимость. С учетом тренда на рост цен на энергоресурсы это обстоятельство повышает привлекательность инвестиций в ВИЭ, которым доставка топлива не требуется. При этом в отличие от ветровых электростанций (ВЭС) и солнечных электростанций (СЭС), чья выработка стохастична по природе и при этом, связана с локальными экологическими проблемами [4-6]. Следует отметить, что гидроэлектростанция (ГЭС) является полностью управляемым и предсказуемым источником при условии качественного прогнозирования местного потребления и имеющихся водных ресурсов.

Несмотря на малый масштаб, такие сети имеют те же особенности, что и объединенные энергосистемы, но в ярко выраженной форме. Так, одной из характерных черт любой энергосистемы является одновременность электропотребления и генерации, а также невозможность накопления этой энергии в значительных объемах. Вследствие этого возникает задача поддержания постоянного баланса мощности между потреблением и генерацией с целью минимизации дефицита мощности и поддержания частоты тока в сети. В ИЭС эта проблема обостряется трудно прогнозируемым резкоизменяющимся потреблением наравне с ограниченной мобильностью имеющихся резервов генерации. Следствием этого становятся колебания частоты в широком диапазоне, а также недоотпуск электроэнергии при ошибках планирования расхода имеющихся ресурсов и неточных краткосрочном и среднесрочном прогнозировании потребления.

В данном исследовании акцент сделан на ИЭС с преобладанием ГЭС. Несмотря на то, что в отличие от прочих устройств ВИЭ гидрогенераторы абсолютно управляемы в краткосрочном интервале, их располагаемая мощность зависит от имеющегося водного ресурса. Это накладывает интегральное ограничение на выработку электроэнергии при планировании режимов на среднесрочном и долгосрочном интервале: при интенсивном расходе воды в одной

части периода планирования, её запас может оказаться недостаточным к его завершению, что приведет к дефициту мощности, вводу резервов дизельной генерации, а также к необходимости поддерживать в работоспособном состоянии такие резервы. В условиях изменяющихся и неопределенных климатических условий, а также невозможности установки систем накопления электроэнергии достаточной мощности и емкости, возникает задача планирования режимов генерации ИЭС с учетом оптимального распределения ресурсов.

Примером такой системы, на базе которой апробируется предлагаемая методика, является ИЭС Горно-Бадахшанской автономной области (ГБАО) Республики Таджикистан. ИЭС ГБАО характеризуется преобладающей долей гидрогенерации малой мощности деривационного типа. Также в системе действует ГЭС с водохранилищем для суточного регулирования. Данная ИЭС испытывает серьезные трудности, связанные с устойчивым дефицитом генерации электроэнергии в зимний период – с ноября по март, а также избытком мощностью в летний период – с мая по сентябрь [7-9]. В летний период ИЭС на 100% обеспечивается энергией от ГЭС, тогда как в зимний период, с увлечением спроса на электроэнергию и понижением уровня воды в реках, требуется задействование резервов дизельных электростанций (ДЭС). Причинами проблем в обеспечении балансовой надежности являются невозможность сезонного регулирования, климатические условия региона и плохая прогнозируемость преобладающей бытовой нагрузки. Применение ДЭС для покрытия небаланса приводит к увеличению себестоимости генерации, ухудшению экологической ситуации данного региона, а также необходимости поддержания этих источников в работоспособном состоянии при том, что большую часть года их работа не требуется. По этой причине ввод новых генерирующих мощностей, требующий значительных капиталовложений, также нежелателен, поскольку коэффициент использования установленной мощности этих электростанций в летний период будет невелик.

Проблема может быть решена, если от принятой на данный момент системы краткосрочного регулирования в течение суток, не позволяющей обеспечить

устойчивое электроснабжение в зимний период, перейти к среднесрочному регулированию в течение недели, учитывая возможности всех генерирующих установок ИЭС. Увеличение горизонта регулирования необходимо для поддержания достаточного запаса гидроресурсов с целью покрытия пикового электропотребления, а также оптимизации резервов мощности и электроэнергии в ИЭС.

В настоящей диссертационной работе разработана модель среднесрочного прогнозирования электропотребления, а также планирования графика генерации электроэнергии с учетом переменного притока воды и температуры окружающей среды. Данная модель является основой для оптимального выбора состава генерирующего оборудования по комплексному критерию: минимум издержек на генерацию электроэнергии при максимуме балансовой надежности ИЭС.

Степень научной разработанности темы исследования. Наибольшее влияние на данную научную область оказали труды следующих учёных: Д.А. Арзамасцева, П.И. Бартоломея, В.З. Манусова, Н.И. Воропая, В.П. Обоскалова, К.В. Сулова, А.А. Мызина, И.И. Надтоки, Т.А. Филипповой, Л.А. Мелентьева, Ю.А. Секретарева, В.М. Горнштейна, М.Г. Тягунова, П.С. Борща, Б.И. Аюева, А.Г. Русиной, М.Ш. Мисриханова и др. Основная часть работ в этом направлении акцентирована на разработке методов прогнозирования электропотребления и генерации крупных энергообъединений. Однако в работах этих авторов в намного меньшей степени уделяется внимание разработке моделей среднесрочного прогнозирования электропотребления и генерации в изолированных энергосистемах с учётом происходящих климатических изменений и использования метеорологических факторов.

Целью исследования является разработка модели среднесрочного прогнозирования электропотребления и генерации с учётом метеорологических факторов в среднесрочной перспективе с целью обеспечения баланса генерации и потребления электроэнергии в изолированных энергосистемах, а также модели

комплексного планирования графика генерации для уменьшения использования дизельных электростанций (ДЭС) за счёт повышения точности прогнозирования и оптимизации регулирования выработки гидроэлектростанций (ГЭС).

Для реализации поставленной цели были сформулированы следующие задачи:

1. Анализ литературы по существующим методам прогнозирования электропотребления и генерация, их достоинств и недостатков. Определение наиболее подходящего метода.

2. Разработка модели среднесрочного прогнозирования электропотребления.

3. Разработка модели среднесрочного прогнозирования притока воды для планирования генерации.

4. Разработка модели среднесрочного прогнозирования генерации гидроэлектростанциями.

5. Разработка модели комплексного планирования графика генерации на среднесрочную перспективу для изолированных энергосистем.

Объектом исследования является изолированная электроэнергетическая система Горно-Бадахшанской автономной области в Республике Таджикистан.

Научная новизна работы заключается в следующем:

1. Впервые для изолированной энергосистемы с высокой долей ГЭС обоснована возможность применения методов машинного обучения для среднесрочного прогнозирования электропотребления и генерации.

2. Разработана новая эффективная модель среднесрочного прогнозирования электропотребления, обеспечивающая повышение точности за счет использования метеорологических данных и кластеризации метеорологических условий.

3. Разработан комплекс авторских моделей среднесрочного прогнозирования притока воды в водохранилища ГЭС на основе ансамблевых

методов машинного обучения для определения объемов воды, достаточных для регулирования генерации в изолированной системе, которые учитывают метеорологические факторы и климатические изменения и отличаются полной автономностью и автоматизацией обучения.

4. Разработана новая модель комплексного планирования графика генерации на среднесрочную перспективу для изолированных энергосистем с высокой долей малых ГЭС деривационного типа на основе адаптивного бустинга линейных регрессий с регуляризацией Тихонова, обеспечивающая близкий к оптимальному выбор состава генерирующего оборудования для обеспечения балансовой надежности.

Теоретическая значимость проведенных исследований заключается в описании эффективного способа применения методов машинного обучения в задачах прогнозирования электропотребления и генерации в изолированных энергосистемах с учётом метеорологических факторов.

Практическая значимость проведенных исследований. На примере реальной изолированной энергосистемы показана принципиальная возможность применения всех разработанных моделей для составления графиков генерации и электропотребления. Результаты практической апробации показали, что возможно использовать предложенные методы для изолированных энергосистем с большой долей гидроэлектростанций.

Результаты, полученные в работе, приняты для использования в энергетических компаниях ОАО «Pamir Energy», который управляет изолированной электроэнергетической системой ГБАО

Методология и методы исследования. Для решения поставленных задач использовались как общенаучные (анализ, синтез, абстрагирование), так и специальные методы математического и компьютерного моделирования электроэнергетических систем, включая проведение вычислительных

экспериментов. Для решения задач прогнозирования были применены методы машинного обучения и алгоритмы обработки данных. К последним относятся системный анализ, математическое и компьютерное моделирование, вычислительный эксперимент, мета-оптимизация. Алгоритмы обработки данных реализованы на языке Python с использованием библиотек: Scikit-Learn и Tensorflow. Визуализация данных выполнено с помощью библиотек: Matplotlib (Python) и Seaborn (Python). Для решения оптимизационных задач применены роевые и эволюционные методы.

Достоверность полученных результатов подтверждается результатами проведенных исследований, представленных на международных конференциях, а также, корректным использованием математического аппарата, сопоставлением полученных результатов и актами внедрения.

Положения, выносимые на защиту:

1. Необходимость применения методов машинного обучения для среднесрочного прогнозирования электропотребления и генерации.
2. Разработанная модель среднесрочного прогнозирования электропотребления для планирования режима загрузки генераторов позволяет обеспечить поддержание баланса генерации и потребления электроэнергии в изолированных энергосистемах.
3. Разработанная модель среднесрочного прогнозирования электропотребления для планирования режима загрузки генераторов позволяет оптимизировать режим работы генераторов с учетом располагаемых гидроресурсов.
4. Разработанная модель среднесрочного прогнозирования притока воды в водохранилище и генерации ГЭС обеспечивает оптимальный выбор состава генерирующего оборудования, определение перспективного запаса воды и оптимальное планирование графика сработки водохранилища.

5. Разработанная модель комплексного планирования графика генерации на среднесрочную перспективу для изолированных энергосистем способна повысить эффективность использования ископаемого топлива на ДЭС, а также снизить выбросы в атмосферу и холостые сбросы на ГЭС.

Апробация результатов. Основные положения данной работы были доложены на научных семинарах кафедры «Автоматизированные электрические системы» УралЭНИН УрФУ, Екатеринбург, в период с 2017 года по 2021 год, а также на нескольких международных конференциях:

- 8th International Conference on Power and Energy Systems Engineering (CPESE 2021), Fukuoka, Japan, 2021;
- 4th International Conference on Electrical Engineering and Green Energy (CEEGE), Munich, Germany, 2021;
- International Conference on Energy Storage Technology and Power Systems (ESPS 2022) Guilin, Chine, 2022.

Личный вклад автора заключается в разработке модели среднесрочного прогнозирования электропотребления, расхода притока воды и генерации гидроэлектростанциями в изолированных энергосистемах, а также модели комплексного планирования графика генерации.

Публикации. По теме диссертации опубликовано 11 научных работ в рецензируемых научных журналах и изданиях, определенных ВАК РФ и Аттестационным советом УрФУ, в том числе 5 работ в журналах, индексируемых в международных реферативных базах Scopus и Web of Science.

Структура и объем работы. Диссертация состоит из введения, 4-х глав, заключения и библиографического списка из 140 наименований. Общий объем работы составляет 119 страниц, включая 40 рисунков и 15 таблиц.

ГЛАВА 1 ОЦЕНКА СОСТОЯНИЯ ЭНЕРГЕТИЧЕСКИХ РЕСУРСОВ ГОРНО-БАДАХШАНСКОЙ АВТОНОМНОЙ ОБЛАСТИ

1.1 Описание объекта исследования

В Республике Таджикистан за выработку, передачу и распределение электроэнергии отвечает Открытая Акционерная Холдинговая Компания (ОАХК) «Барки Точик». Единственной областью страны, электроснабжением которого занимается частная энергокомпания «Pamir Energy» является ГБАО. Данная компания в течение 25 лет на основе Концессионного соглашения начиная с 2002 года осуществляет электроснабжение региона. Необходимо отметить, что электроснабжение ГБАО осуществляется изолированно, т.е. без связи с основной энергосистемой Республики.

В настоящее время в управлении «Pamir Energy» находятся генерация, установленная мощность которой достигает 44 МВт, а именно: 11 ГЭС (Таблица 1.1) из которых более крупные - ГЭС «Памир-1» и «Хорог» (Рисунок 1.1), одна солнечная электростанция (СЭС), а также несколько ДЭС малой мощности. Следует отметить, что в ГБАО также функционируют более десятка малых ГЭС, которые не подключены в общую сеть, работают в разбросанных по области и удаленных горных поселках и не находятся на балансе «Pamir Energy». По этой причине работа таких станций в диссертационной работе не рассматривалась.

Таблица 1.1 – Сведения о функционирующих ГЭС на территории ГБАО, находящихся на балансе «Pamir Energy»

№ п/п	Наименование ГЭС	Установленная мощность МВт	Год ввода в эксплуатации	Кол-во гидроагрегатов	Расчетный напор (м)	Расход воды через агрегат м ³ /с	Место расположение
1	ГЭС Памир-1	28	1994	4	79,6	10,1	Шугнан
2	ГЭС Хорог	9	1970	5	59	3,55	Город Хорог
3	ГЭС Намангут	2,5	1974	2	36	3,5	Ишкашим

4	ГЭС Таджикистан	1,5	2018	2	31	3,2	Мургаб
---	--------------------	-----	------	---	----	-----	--------

Продолжение таблицы 1.1

5	ГЭС Ванч	1,2	1968	2	21,5	3,5	Ванч
6	ГЭС Шуджанд	0,832	1969	2	10	5,5	Рушанский район
7	ГЭС Техарв	0,36	1995	1	110	0,59	Ванч
8	ГЭС Андарбак	0,3	1999	1	23	1,8	Ванч
9	ГЭС Калаи- Хумб	0,208	1959	2	10,8	1,55	Дарвоз
10	ГЭС Сипондж	0,16	1992	2	130	0,31	Рушан
11	ГЭС Савноб	0,08	1984	1	72	0,1	Рушан

1.2 Мургабские районные электрической сети (РЭС)

Среди районов ГБАО особо выделяется Мургабский, расположенный в его восточной части (Рисунок 1.1). При площади 38442,2 км² все население этого района (почти 17 000 [10]) живет в 20 населённых пунктах, разнесенных на 50 - 150 км друг от друга. Из-за особенностей рельефа и взаимной удаленности, центры потребления при развитии энергосистемы ИЭС ГБАО в советское время было решено не подключать к системе централизованного электроснабжения в силу экономической нецелесообразности. Альтернативой стали распределенные ДЭС малой мощности, обеспечившие покрытие более 70% потребления региона. Затраты на реализацию такого решения также были значительны из-за высокой стоимости дизельного топлива для этих станций с учетом доставки из центральных регионов страны и низкого КПД установок. При планировании развития практически не учитывалась экологичность генерации, хотя вред окружающей среде от дизельных генераторов по сегодняшним меркам значителен [11-13]. Оптимизация затрат и особенности местности обусловили низкую надежность электроснабжения потребителей. Ограниченное число ДЭС при отсутствии внешних резервных источников не позволяло обеспечить резервирование по мощности в случае отказа установок. Сложности с транспортным сообщением

приводили к увеличению продолжительности восстановительных работ, а также развитию проблем с резервированием по энергии – обеспечением ДЭС топливом. При эксплуатации постоянные колебания частоты и мощности, характерные для изолированных энергосистем [14], приводили к ускоренному износу генераторов, росту вероятности отказов, что в еще большей степени влияло на надежность электроснабжения потребителей Мургабского района [15]. В этих условиях экономическое развитие области проходило крайне медленно: население преимущественно занималось натуральным хозяйством, промышленность находилась в зачаточном состоянии. Ситуация осложнилась после распада СССР, когда в условиях ослабления контроля, а также экономических и политических катаклизмов тех лет в эксплуатации не осталось ни одной ДЭС. Наглядным результатом этих изменений стало то, что на 2002 год только 13% мургабских семей имели доступ к электроэнергии при постоянных перебоях электроснабжения.

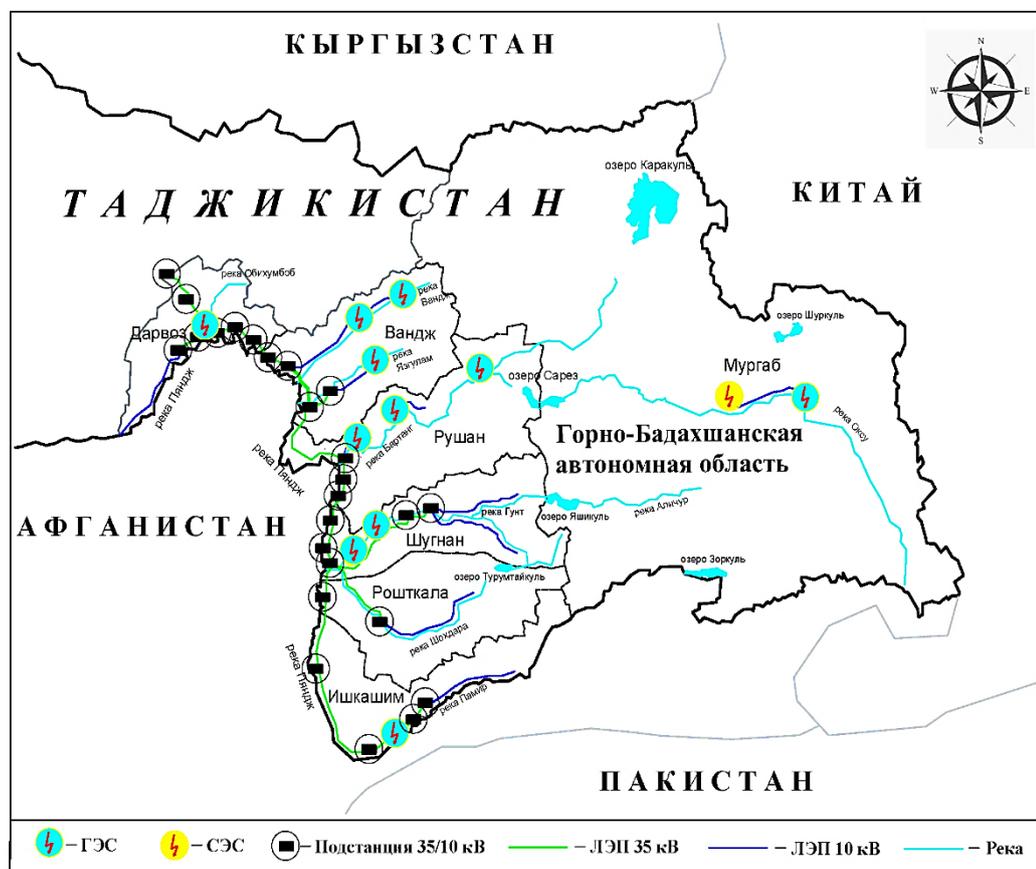


Рисунок 1.1 – Географическое расположение ИЭС ГБАО

Климат в районе сухой и крайне холодный. Температура в Мургабском районе с ноября по март регулярно опускается до -50 градусов по Цельсию. Температура в летний период на высотах $3500\text{--}4200$ метров колеблется в пределах $+11\dots+12$ градусов. В июле возможны ночные заморозки. В зимний же период реки замерзают, вследствие чего рабочая мощность ГЭС снижается вплоть до их вынужденной остановки. Например, самая большая река в районе малосточная река «Ак-су» может замерзнуть на 80% . По этим причинам гидроэнергетика, несмотря на свою экологичность, не может круглогодично быть доступным источником электрической энергии в районе. Для решения этой проблемы USAID поддержало пилотный проект в 2020 г., по строительству СЭС, которая могла бы обеспечивать население Мургаба электричеством в зимний период [16].

На территории Мургабского района на сегодняшний день работает много распределенных по территории малых ГЭС, которые экономически не выгодно соединить их в единую сеть. Начиная с 2020 г. ситуация меняется в положительную сторону, введена в эксплуатацию одна ГЭС ($1,5$ МВт) с водохранилищем с запасом воды на одни сутки работы с установленной мощностью и одну СЭС (220 кВт), функционирующих как гибридный энергокомплекс для обеспечения электроснабжения города Мургаба. В будущем планируется подключить другие соседние генерирующие станции в этом энергокомплексе, образовав таким образом мини-автономную изолированную энергосистему.

На сегодняшний день район Мургаба работает изолированно от основной энергосистемы ГБАО, так как строительство протяженных ЛЭП до местных потребителей малой мощности по-прежнему экономически нецелесообразно. Кроме того, наличие таких связей в принципе не может быть альтернативой собственной распределенной генерации, из-за их потенциально высокой аварийности и больших интервалах восстановительных и ремонтных работ. Это привело бы к перерывам питания и общему снижению надежности, что в свою очередь стало бы основанием для ввода и поддержания в работоспособном состоянии резервов мощности ДЭС в удаленных населенных пунктах.

Приведенные особенности Мургабского района характерны для ГБАО и демонстрируют основную проблематику работы энергосистемы в данном регионе в целом. Разница, тем не менее, заключается в том, что в западной части ГБАО, из-за местной сравнительно высокой плотности населения, было организовано централизованное электроснабжение, что актуализировало задачи прогнозирования потребления и планирования генерации. Разрозненность, автономность и малые мощности нагрузок Мургабского района не представляют интерес с точки зрения решения этих задач, поэтому прогнозирование потребления и генерации для данного района в работе не проводится.

1.3 Потенциал альтернативных источников энергии ГБАО

1.3.1 Гидроэнергетические и водные ресурсы

ГБАО благодаря своему географическому расположению, является обладателем большого гидроэнергетического потенциала. Гидроэнергетика, как было отмечено выше, круглогодично является основным источником электроэнергии для ИЭС ГБАО. Несмотря на это, на сегодняшний день имеющиеся ресурсы используются только на 6-7% [17].

Выработка электроэнергии малыми ГЭС ГБАО при полном использовании имеющегося потенциала многократно превышает внутренние потребности региона даже с учетом отдаленной перспективы. С целью поддержания и развития малой гидроэнергетики, правительством страны был принят ряд законов и постановлений, в том числе программы краткосрочного, среднесрочного, и долгосрочного развития малой гидроэнергетики и нетрадиционных ВИЭ [18-20]. Данная программа создает условия для сооружения малых ГЭС мощностью от 1–30000 кВт, что станет хорошей основой для развития гидроэнергетики особенно в горных селах региона. Так, в регионе имеется возможность обеспечить свои отдаленные горные и труднодоступные села электроэнергией примерно на 60–75%, а в некоторых случаях до 100% за счет малых рек [21]. Также, в перспективе

планируется строительство ГЭС «Себзор» мощностью 10 МВт и ГЭС «Санобод» мощностью более 125 МВт в Рошткалинском и Рушанском районах соответственно. Таким образом, малая гидроэнергетика способна быть один из прибыльных направлений для формирования энергетики региона.

Строительство мини и малых ГЭС в условиях Республики Таджикистан имеет удельную стоимость 2000–3000 дол США за 1 кВт установленной мощности. При этом, срок окупаемости этих станций составляет 3-8 лет [22].

При таких ценах в регионе, любая малая ГЭС, построенная даже в самых сложных условиях климата и рельефа, оказывается эффективнее альтернативной тепловой генерации и в то же время благодаря управляемости превосходит другие возобновляемые источники.

Таблица 1.2 – Гидроэнергетические ресурсы ГБАО по категориям

Регион	Промышленные запасы		Потенциальные ресурсы					
	МВт	ТВт/ч	Крупных рек		Притоков L>10км		Притоков L<10км	
			МВт	ТВт/ч	МВт	ТВт/ч	МВт	ТВт/ч
Горно-Бадахшанская – Автономная область	5884	51,54	6990	61,23	2555	22,38	3713	32,53

Таблица 1.3 - Потенциальные запасы гидроэнергоресурсов ГБАО

Бассейны рек	Среднегодовая мощность, МВт.	Среднегодовая энергия, ТВт·ч.	Доля в общем объеме, %
Пяндж	9420	82,52	55,8
Бартанг	2969	26,01	17,7
Гунт	2260	19,8	13,5
Ванч	1191	10,34	7
Язгулем	845	7,4	5,01
Оз.-Кара-Куль	103	0,9	0,6
Итого	16788	147,06	100

Таблица 1.4 – Технические и экономические гидроэнергетические ресурсы малых водотоков ГБАО

№п/п	Районы	Ко-во створов	Технический потенциал		Экономический потенциал	
			МВт	ТВт ч	МВт	ТВт ч
1	Дарвазский	8	89,39	0,783	35,7	0,313
2	Ванчский	7	47,91	0,420	19,16	0,168
3	Рушанский	13	76,08	0,667	30,43	0,267
4	Шугнанский	14	76,03	0,667	30,43	0,267
5	Рошткалинский	12	30,59	0,268	12,23	0,104
6	Ишкашимский	11	42,24	0,370	16,89	0,148
7	Мургабский	8	22,33	0,196	8,93	0,078
8	Всего по ГБАО	73	384,57	3,372	153,83	1,349

На ряду с гидроэнергетическими ресурсами и другие виды ВИЭ распределены равномерно по территории региону (Рисунок 1.2).

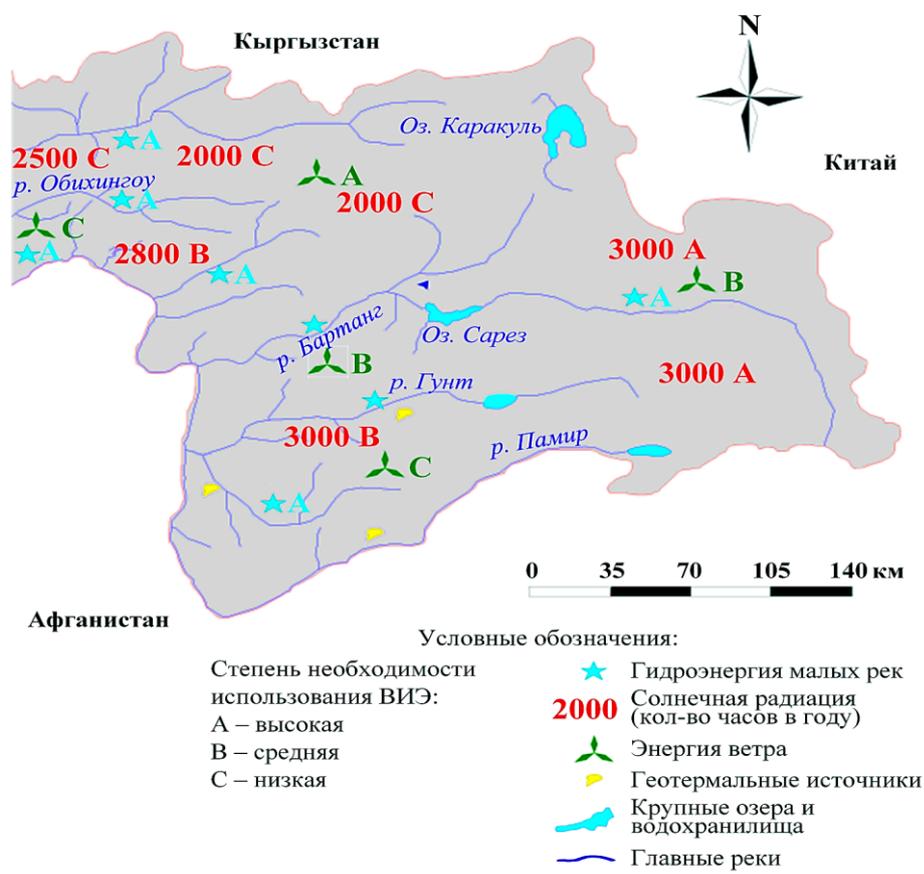


Рисунок 1.2 – Потенциал альтернативных источников энергии ГБАО

1.3.2 Солнечные ресурсы

ГБАО имеет благоприятные условия для использования солнечной энергии и считается одним из наиболее перспективных регионов центральной Азии с точки зрения инвестиций в строительство солнечной генерации.

В целях обеспечения устойчивого и экологически безопасного развития энергетики региона, солнечная энергия может играть важную роль в этом направлении. Суммарная солнечная радиация при ясном небе достигает 700–800 Вт/м² или 7,500–8,000 МДж/м². Интенсивность прямой солнечной радиации изменяется от 1,30 до 1,7 кал/см²/мин [23]. Эти параметры намного выше в горных районах: особенно в описанной выше Восточной части ГБАО, где население имеет ограниченную возможность использования гидроэнергетических ресурсов, а доставка топлива затруднена и экономически неэффективна. Тем не менее, следует отметить, что в этих районах также наблюдается и наименьшая продолжительность солнечного сияния: характерны закрытый рельеф и пасмурная погода (ледник Федченко - 2218 часов в год). Наибольшая продолжительность солнечного сияния (более 3000 часов в год) наблюдается в Северо-Восточной части ГБАО (озеро Каракуль - 3280 часов) [24]. Эта территория является наиболее подходящей для использования солнечной энергии.

Описанные обстоятельства создают хорошие предпосылки для масштабного вовлечения ВИЭ в топливно-энергетический баланс региона в целях обеспечения устойчивого экологически безопасного развития энергетики региона. При этом возникает перспективная задача развития солнечной генерации не только для покрытия местного потребления, но и поиска оптимальных путей максимальной реализации потенциала по выработке электроэнергии и её передачи в соседние регионы с учетом указанных выше трудностей. Исследованию этой проблемы будут посвящены будущие работы.

Необходимая информация об оценке потенциальных ресурсов солнечной энергии на территории ГБАО приведена в таблице 1.5 – 1.6.

Таблица 1.5 – Число дней без солнца в ГБАО

Наименования станция	Месяцы												За год
	1	2	3	4	5	6	7	8	9	10	11	12	
Кара-Куль	2	2	2	0	0	0	0	0	0	0	1	2	9
Ледник Федченко	12	12	11	6	3	1	1	0	0	5	10	12	73
Хабурабад	6	5	5	3	2	0	0	0	0	1	6	8	36
Мургаб	4	3	2	1	0	0	0	0	0	1	2	3	16
Хорог	7	4	3	1	0	1	0	0	0	1	2	6	25

Таблица 1.6 - Суммарная месячная солнечная радиация в административном центре ГБАО Вт/м²

Станция	Месяцы											
	1	2	3	4	5	6	7	8	9	10	11	12
Хорог	96	137	187	320	304	350	340	305	258	172	114	86

1.3.3 Ветровые ресурсы

В ГБАО система метеорологических наблюдений не организована должным образом, статистические данные ограничены по времени и не покрывают всю территорию [25]. По этой причине в регионе реальный потенциал энергии ветра только предстоит оценить и в данный момент исследователями выполняются научные и экспериментальные работы в этом направлении.

Исходя из имеющихся данных потенциал развития ветровой генерации для различных районов ГБАО неравномерен. Так, например, среднегодовая скорость ветра на леднике Федченко составляет 6,0 м/с, а в типичных местах как озеро Каракуль – 3,0-3,7 м/с. В сельских районах таких как Рушан, Хорог, Мургаб, и Ишкашим – 2,0-2,7 м/с. В целом, в большинстве остальных регионов страны средняя скорость ветра незначителен и варьируется в пределах от 0,9-4,8 м/с [26].

В данный момент ветрогенерация не может конкурировать с действующей гидроэнергетикой и солнечной генерации. На неравномерность и территориальную распределенность точек потребления накладывается низкая стоимость

электроэнергии. В ГБАО, согласно [27], тарифы на электроэнергию составляют – 0,26 сомони = 2,5 цент/ кВт·ч. При этом в мире минимальная стоимость электроэнергии, при которой целесообразно использование ветрогенерации составляет 4 цент/кВт·ч. Минимальная стоимость электроэнергии, при которой экономически выгодно строительство и тарифы на ветроэнергетику в мире не ниже 4 цент/кВт·ч. Исходя из этого, в данный момент ветроэнергетика в ГБАО не может конкурировать с другими видами генерации, но может быть востребована локально в первую очередь в социально-бытовой сфере.

Подробная информация об оценке потенциальных ресурсов ветровой энергии на территории ГБАО приведена в таблице 1.7.

Таблица 1.7 -Среднегодовая скорость ветра по ГБАО

Станция	Высота над уровнем моря, м	Скорость ветра, м/с
Калай-Хумб	1288	1,3
Мургаб	3576	2,3
Рушан	1981	2,1
Ишкашим	2524	2
Хорог	2075	2,3
Хумгари	1736	2,7
Джавашангоз	3410	2,3
Булункуль	3744	1,5
Озера Каракуль	3930	3
Ледник Федченко	4169	6
Перевал Хабурабад	3347	4,8

1.4 Влияние ДЭС в ИЭС ГБАО

За последние двадцать лет (с декабря 2002 г.), после того как сеть электроснабжения ГБАО перешла от компании «Барки Точик» к частной компании «Pamir Energy» произошло много изменений на мировом рынке из-за высокой

стоимости нефтепродуктов. В целом, в Республике Таджикистана отсутствуют месторождения газа и нефти, а в ГБАО отсутствуют и месторождения угля. Все, что связано с углеродным топливом в регионе, импортируется и получается очень дорогим из-за труднодоступности региона.

В общей структуре генерирующих мощностей ИЭС ГБАО доля ДЭС в конце своего полного распада СССР (26. 12. 1991 г.) составляла более 60% (Таблица 1.2). В зимний период (с ноября по марту) из-за резкого снижения гарантированных мощностей на ГЭС по причине уменьшения притока воды, показатель участия ДЭС в покрытии электрической нагрузки потребителей колебался от 75 до 80%, при параллельной работе ГЭС и ДЭС в электрической сети.

Таблица 1.8 – Установленные и располагаемые мощности ДЭС по Горно-Бадахшанским электрическим сетям (ЭС) на 26. 12. 1991 г.

№п/п	Наименование станции по районам области	Место установки	Высота над уровнем моря, м	Установленная мощность, кВт	Располагаемая мощность, кВт
	1.Мургабский РЭС				
1.	ДЭС Мургаб	Мургаб	3592	3200	2000
2.	ДЭС Кизил Чигарчи	Каракуль	3785	155	101
3.	ДЭС Ленин -Жолу	Рангкуль	3514	120	78
4.	ДЭС Киров	Шаймак	3650	60	48
5.	ДЭС Киров	Тохтамиш	3635	170	122
	Всего			3675	2325
	2.Ишкашимский РЭС				
1.	Σ ДЭС	Ишкашим	2900	2500	2000
	3. Шугнанский РЭС				
1.	ДЭС ПАЭС	Хорог	2100	6600	5500
2.	ДЭС Шугнан	Хорог	2100	6400	6280
3.	ДЭС Буни	Буни	2050	2400	2280
	Всего			15400	12980
	4. Рушанский РЭС				
1.	ДЭС Шуджанд	Шуджанд	1990	2400	2280

Продолжение таблицы 1.8

2.	ДЭС Интернационал	Спондж	2400	225	225
3.	ДЭС Савноб	Савноб	2700	30	20
4.	ДЭС Басид	Басид	2750	74	74
	Всего			2899	2359
5. Ванчский РЭС					
1.	ДЭС Ванч	Ванч	1750	2400	2067
6. Калаи - хумбский РЭС					
1.	ДЭС Калаи-Хумб	Калаи-Хумб	1260	800	720

Как отмечалось выше, в условиях роста цен на энергоресурсы и после перехода к рыночным отношениям использование ДЭС в ГБАО в отсутствие государственных дотаций стало неэкономичным, что сравнительно быстро привело к их остановке и выводу из эксплуатации. Тем не менее, суровые погодные условия характерные для зимнего периода, которые при прекращении электроснабжения поставят население региона на грань выживания, вынуждают энергетические компании страны искать технически обоснованные альтернативные пути преодоления зависимости от дорогих импортных нефтепродуктов. Для Республики Таджикистан и в особенности для ГБАО, сосредоточившей огромные гидроэнергетические ресурсы страны, ответом стали инвестиции в развитие гидроэнергетики – строительство малых, средних и больших ГЭС.

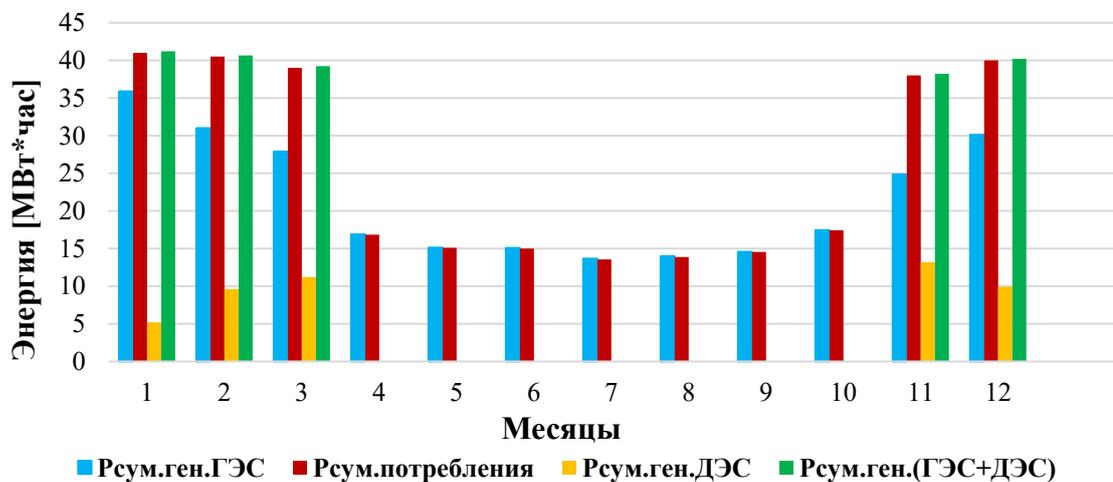


Рисунок 1.3 – График генерации и потребления за 2020 г.

С помощью привлеченных многомиллионных иностранных грантов компании «Pamir Energy» удалось преуспеть в этом вопросе, что позволило дополнительно снизить зависимость от ввозимого углеродного топлива и в весенне-летний период (с апреля по октябрь включительно) перейти на выработку электроэнергии исключительно за счет ГЭС. На рисунке 1.3 показан годовой профиль потребляемой электроэнергии по месяцам.

Как можно видеть, в зимний период уровень нагрузки в области резко возрастает. В эти месяцы удовлетворить спрос на электроэнергию только за счет ГЭС не удастся, из-за снижения их рабочей мощности при низких температурах. В таких случаях для покрытия нагрузки используются ДЭС.

На рисунке 1.4. приведена однолинейная схема энергосистемы ГБАО. В этой схеме ДЭС используются только в качестве резервных источников.

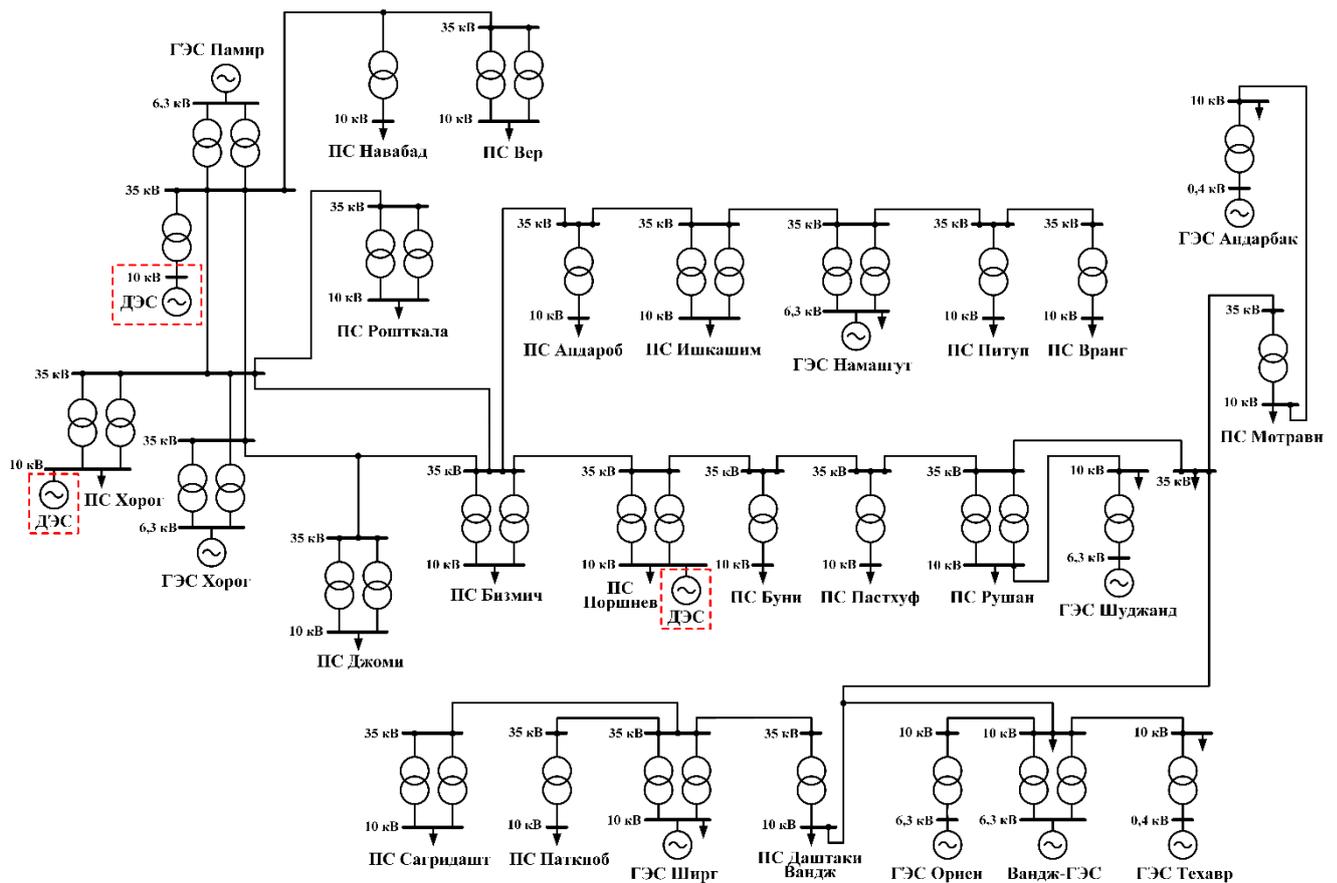


Рисунок 1.4 – Однолинейная схема рассматриваемой изолированной энергосистемы ГБАО

Выводы по главе 1

В рамках первой главы диссертационного исследования основное внимание было уделено оценке состояния энергетических ресурсов Горно-Бадахшанской автономной области. Показано, что регион богат возобновляемыми энергетическими ресурсами, особенно гидроэнергетикой.

В настоящее время в ГБАО нет промышленного использования потенциала солнечной энергии, хотя климатическое расположение региона достаточно благоприятно для использования солнечной энергии. Широкое использование солнечной энергии оказало бы положительное влияние на энергетическую и экономическую ситуацию в регионе. Что касается энергии ветра, то в настоящее время она экономически неконкурентоспособна с действующей гидроэнергетикой и солнечной генерацией.

ГБАО испытывает серьезные трудности со устойчивым электроснабжением зимой, когда уровень воды в реках падает. Необходимость увеличения доли микро-, мини- и малых ГЭС и сокращения доли существующих ДЭС в зимний период, поскольку в регионе практически нет месторождений угля, нефти и газа. Из-за сложного расположения региона и транспортировки привозные нефтепродукты окажутся очень дорогими.

ГЛАВА 2 АНАЛИЗ МЕТОДОВ СРЕДНЕСРОЧНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ И ГЕНЕРАЦИИ

2.1 Классификация методов прогнозирования электропотребления и генерации

Прогнозирование электропотребления и генерации является базой для принятия управленческих решений в электроэнергетических системах. Точность прогнозов напрямую влияет на эффективность планирования и управления [28]. Прогнозирование электропотребления и генерации являются важным условием обеспечения бесперебойной и эффективной работы энергосистем, а также важной основой для создания новых цифровых и интеллектуальных энергосистем [29]. Учитывая, что на электропотребление энергосистемы влияют различные факторы (например, изменение климата и окружающая среда, метеорологические факторы), электропотребление энергосистемы имеет сильную случайность и изменчивость, будучи при этом отчасти периодической величиной. Задача прогнозирования генерации считается одной из самых сложных задач в области интеллектуального анализа данных, требующая комплексного анализа больших объемов данных с учетом влияния на электропотребление множественных взаимосвязей и динамических процессов [30].

С точки зрения горизонтов прогнозирования, прогнозы подразделяются на четыре различные категории:

- **Оперативное прогнозирование:** прогноз делается на несколько минут вперед, и прогнозируемые значения передаются блоку управления для отправки в режиме реального времени. Это прогнозирование используется для быстрого реагирования на внутрисуточные колебания спроса на электроэнергию [31].
- **Краткосрочное прогнозирование:** прогноз делается от нескольких часов до нескольких дней вперед, и результаты используются для широкого спектра решений, связанных с технико-экономической оптимизацией функционирования электроэнергетических систем и обязательствами по качественному и надежному

электроснабжению потребителей электрической энергией [32]. В настоящее время значимость оперативного и краткосрочного прогнозирования возрастает по причинам, связанным с усложнением систем транспорта и распределения электроэнергии, возрастающим требованиям по качеству и надежности электроснабжения, а также появлением нового оборудования с автоматизированными системами управления.

- Среднесрочное прогнозирование: в этом случае прогноз делается от нескольких дней до нескольких месяцев вперед. Полученные прогнозы дают информацию о недельных колебаниях, и эта информация в основном используется для планирования технического обслуживания сети, установки цен на электроэнергию и согласования механизма распределения энергии и т. д. [33].

- Долгосрочное прогнозирование: прогноз делается от нескольких месяцев до нескольких лет, и данная информация обычно используется для оценки нагрузочной способности и анализа необходимости развития электросетевой инфраструктуры [34].

Большинство методов прогнозирования электропотребления и генерации посвящены краткосрочному прогнозированию (от нескольких минут до 24 часов), а не среднесрочному (от нескольких дней до нескольких месяцев). Как правило, методы прогнозирования электропотребления и генерации можно разделить на две широкие категории [35]: статистические методы и методы на основе машинного обучения, хотя граница между ними становится все более и более размытой. Далее методы первой группы обозначаются A1, A2, A3, A4 (Статистические) и методы второй группы B1, B2, B3, B4 (основанные на машинном обучении).

Традиционные статистические методы включают методы анализа временных рядов [36], метод фильтра Калмана [37], метод экспоненциального сглаживания [38] и так далее. Однако точность традиционных методов прогнозирования может быть недостаточно высокой на среднесрочную перспективу из-за нелинейных характеристик электропотребления для прогнозирования процессов с высокой степенью волатильности. Последнее особенно характерно для ИЭС из-за их небольшого размера относительно крупных региональных и объединенных

энергосистем. Методы прогнозирования, основанные на машинном обучении, включают нечеткую логику, классические искусственные нейронные сети (ANN) [39], метод опорных векторов (support vector machine, SVM) [40], нейронные сети с памятью [41] и ансамблевых методов [42]. Их можно рассматривать как нетрадиционную или современную методологию в задачах прогнозирования электропотребления и генерация. По сравнению с традиционными методами точность прогнозирования методов машинного обучения во многих случаях существенно выше, но оптимизация параметров каждой модели затруднена, что влияет на эффективность прогнозирования.

2.2 Математические методы, используемые для среднесрочного прогнозирования электропотребления и генерации

2.2.1 Статистические методы

Статистические методы строят математическую или статистическую модель электропотребления (генерации) путем изучения качественных взаимосвязей между электропотреблением (генерацией) и факторами, влияющими на нагрузку. Таким образом, способ не предлагает пользователю интуитивного понимания. Этот тип моделей включает в себя модели регрессии (линейной и нелинейной), авторегрессионные модели (ARIMA, ARIMAX, GARCH, ARDLM и др.), модели с экспоненциальным сглаживанием, модели максимального правдоподобия. Предполагаемые параметры модели оцениваются на основе исторических данных, а адекватность модели проверяется путем анализа остатков, т.е. ошибок прогноза. Концепция этого метода заключается в установлении нелинейных взаимосвязей между входными данными (предыдущими значениями электропотребления (генерации) и факторами, влияющими на электропотребление (генерацию)) и выходными данными (прогнозируемым электропотреблением (генерацией)) путем их выражения по явной формуле. Методы прогнозирования временных рядов постоянно развиваются, все чаще представляя собой количественный подход с

прошлыми данными в качестве основы для прогнозирования [43, 44], примеры приведены ниже.

Простые статистические модели (A1). Преимущества этих моделей прогнозирования заключаются в простоте, гибкости и согласованности их анализа и проектирования [45, 46]. При использовании моделей линейной регрессии результат может быть получен наиболее быстро по сравнению с использованием других моделей. Кроме того, возможность проанализировать все промежуточные вычисления также является существенным преимуществом. Основным недостатком линейных моделей регрессии является трудность определения нелинейных функциональных зависимостей [47].

Авторегрессионные модели и методы (A2). Преимущества этого класса моделей – простота и прозрачность моделирования. В настоящее время этот класс моделей является одним из самых популярных. Недостатками данного класса моделей являются: большое количество параметров модели, не всегда однозначно определяемых, низкий уровень адаптивности, а также линейность, что ведёт к неспособности моделировать нелинейные процессы, которые часто встречаются на практике. Модель авторегрессионной скользящей средней (ARMA), авторегрессионная модель и модель авторегрессионной интегрированной скользящей средней (ARIMA) применяются наиболее часто [48, 49].

Модели экспоненциального сглаживания (A3). Главные достоинства заключаются в простых вычислениях и гибкости описания разных изменений. В [50] на основе анализа результатов подобных моделей сделан вывод, что уникальную структуру энергии и спроса, относящуюся к быстрорастущим районам, было трудно проанализировать и предсказать прямым применением метода Винтера. Экспоненциальное сглаживание было дополнено анализом спектра мощности и адаптивным авторегрессионным моделированием в гибридном подходе [51]. В работе Филда и Хилла [52] было показано, что метод удаления тренда, основанный на оптимальном сглаживании, выгодно отличается от традиционных методов прогнозирования нагрузки. Данный метод позволяет получать оценки параметров тренда, описывающих не средний уровень явления, а

тенденцию, сложившуюся к моменту последнего наблюдения. Для данного метода очень важно выбрать параметры сглаживания и начальные условия. В то же время такие модели, как правило, не подходят для долгосрочного прогнозирования.

Метод максимального правдоподобия (А4) или метод наибольшего правдоподобия (ML, MLE – англ. maximum likelihood estimation) – это метод оценивания неизвестного параметра путём максимизации функции правдоподобия [53]. Основан на предположении о том, что вся информация о статистической выборке содержится в функции правдоподобия. Для фиксированного набора данных и базовой вероятностной модели, используя метод максимального правдоподобия, мы получим значения параметров модели, которые делают данные «более близкими» к реальным. В [54] для прогнозирования гидрологических параметров осадков, уровней сброса воды, стока осадков и т.д. использовалась модель робастной регрессии максимального правдоподобия (MLRM). Оценка максимального правдоподобия даёт уникальный и простой способ определить решения в случае нормального распределения. Основной недостаток метода максимального правдоподобия – трудность вычисления оценок, связанных с решением уравнений правдоподобия, чаще всего не линейных.

Иногда статистическая модель смешивается с другими методами для формирования гибридной модели, которая обладает большей точностью по сравнению со стандартной статистической моделью или даже обычным ANN. Некоторые примеры таких работ включают гибридизацию модели ARIMA с вейвлет-разложением, усреднение байесовской модели и методы прогнозирования Грея [55-57]. Аналогичным образом, подход Байесовского вывода используется для определения параметров распределения Вейбулла (вероятностное прогнозирование) для прогнозирования скорости ветра [58].

Все приведенные статистические методы в целом проще в реализации, чем другие подходы, экономичны и требуют меньшей вычислительной мощности. Используя эти методы, исследователи смогли получить приемлемые результаты для коротких временных интервалов до 48 часов, но прогнозирование становится очень неустойчивым для более длительных периодов времени [55, 59, 60]. Другим

недостатком этих моделей является то, что для преобразования данных в стандартную форму требуется предварительная обработка данных временных рядов. В таких приложениях используются препроцессоры, такие как независимый компонентный анализ (ICA) и вейвлет-преобразование [61, 62]. Иногда модель временных рядов комбинируется с моделью физического вывода, которая распознает физические закономерности данных о ветре [60]. Гибрид этих статистических моделей с интеллектуальными методами также популярен для преодоления этих проблем [63].

2.2.2 Методы прогнозирования, основанные на машинном обучении

Одним из наиболее эффективных методов для решения задач прогнозирования являются методы машинного обучения, способные выявлять более сложные зависимости, чем более простые регрессионные или авторегрессионные модели. Совершенствование математических методов машинного обучения и увеличение вычислительных мощностей привели к впечатляющим результатам сложных моделей машинного обучения во многих областях. Однако при построении прогностических моделей, особенно на базе машинного обучения, как правило, первый приоритет отдается повышению их точности. Вопросы робастности и обобщающей способности (или отсутствия переобучения) учитываются, но отодвигаются на второй план. Для подтверждения того, что модель не переобучена и нет подгонки модели под выборку, принято использовать разбиение выборки на обучающую, валидационную и тестовую, и/или применять технику кросс-валидации. Если показатели точности модели на обучающей, валидационной и тестовой выборках близки, то делается вывод о том, что модель не переобучена и может применяться в будущем на новых данных, на которых она не обучалась. Этот тип моделей включает в себя: метод опорных векторов (Support Vector Machine — SVM), случайный лес (Random Forest), бустинг (Adaptive Boosting — AdaBoost), метод k-ближайших соседей (k-nearest neighbors algorithm — kNN), и искусственные нейронные сети (ANN). Достоинствами этих

моделей являются простота и однородность их анализа и проектирования. Недостатком является невозможность моделировать процессы на множество шагов вперед. долгой памятью.

Метод опорных векторов (SVM), (B1). Вапник был первым, кто представил SVM; это новый мощный метод машинного обучения, основанный на теории статистического обучения (SLT), который анализирует данные и распознает закономерности, используемые для классификации и регрессионного анализа. Преимуществом метода является возможность учета большого количества дополнительных факторов, влияющих на качество прогнозирования. Они сочетают управление обобщением с методом устранения проклятия размерности [64]. В [65] доказано, что информация о температуре и других параметров климата не очень полезна для среднесрочного прогнозирования электропотребления и внедрение прогнозирования временных рядов может улучшить результаты. В работе [66] модифицированная функция риска обычных машин с опорными векторами путем более строгого наказания за нечувствительные ошибки, чем за удаленные нечувствительные ошибки, они назвали этот метод *C*-восходящей машиной опорных векторов. Они пришли к выводу с помощью теста, что машины с *C*-восходящими опорными векторами с фактически упорядоченными выборочными данными последовательно прогнозируют лучше, чем стандартные машины с опорными векторами. Недостатками метода является медленное обучение и повышенные требования к программно-аппаратным ресурсам.

Ансамблевые методы (Random Forest, AdaBoost), (B2). Основными преимуществами моделей на основе регрессионных деревьев решений являются высокая скорость обучения и статистическая обоснованность процесса обучения, масштабируемость. Кроме того, есть возможность использовать категориальные факторы. Всё это позволяет быстро обрабатывать большие объемы данных. К недостаткам этих моделей можно отнести неоднозначность алгоритма построения дерева, склонность деревьев к переобучению и невозможность интерпретации моделей при наличии в них большого числа узлов, так как в этом случае правил становится слишком много, чтобы их можно было проанализировать, [67].

Алгоритм k-ближайших соседей (kNN), (B3). Одним из наиболее распространенных алгоритмов Машинного обучения является Алгоритм k-ближайших соседей (k-Nearest Neighbor, kNN), популярность которого обусловлена целым рядом неоспоримых достоинств, а именно простоте реализации и прозрачной процедуре работы. Это приводит к высокой интерпретируемости результатов kNN, в отличие от многих других методов машинного обучения, которые создают либо слабо интерпретируемые модели, либо не интерпретируемые модели, такие, например, как ансамбли деревьев решений, нейронные сети, метод опорных векторов и другие [68]. Недостатком kNN является необходимость выбора гипер-параметров, в первую очередь числа соседей k , а также метрики расстояний (distance function).

Искусственные нейронные сети (ANN), (B4). Последние достижения в области машинного обучения (ML) и технологий анализа данных (DA) предоставили возможность более эффективно использовать большое количество данных. В настоящее время наибольшей популярностью пользуются искусственные нейронные сети (ANN). Они позволяют строить сколь угодно сложные модели для выявления сложных зависимостей за счет масштабирования: увеличения числа слоев модели, применения различных элементов обработки данных, таких как простые сумматоры (нейроны), ячейки с внутренней памятью, сверточные локальные фильтры, ячейки множественного выхода; комбинирования различных видов ячеек и слоев [69].

Главным преимуществом нейросетевых моделей является нелинейность, то есть умение устанавливать нелинейную зависимость между будущими и прошлыми значениями. Другие важные преимущества – адаптивность и масштабируемость (параллельная структура искусственной нейронной сети ускоряет вычисления с помощью применения многоядерных графических процессоров). В то же время есть и недостатки: отсутствие прозрачности моделирования, сложность выбора архитектуры сети, высокие требования к объему и качеству обучающей выборки, сложность выбора алгоритма обучения и ресурсоемкость обучения, необходимость настройки большого числа параметров

не самой сети, а алгоритмов ее обучения. Кроме того, их недостаток состоит в «перетренировке сети» (overfitting) и долгом времени обучения в случае сложных архитектур и больших объемов данных (с другой стороны, современные достижения нейронных сетей чаще всего связаны именно с их обучением на огромных наборах данных и увеличении вычислительной сложности моделей) [70].

Следует отметить, что данная выше классификация представляет лишь основные и часто используемые модели прогнозирования, однако в настоящее время таковых насчитывается гораздо больше, особенно с учетом гибридных моделей.

Использование ANN для прогнозирования скорости ветра и солнечной инсоляции все еще можно увидеть в совсем недавних работах [71], обычно в сочетании с другими передовыми методами, такими как вейвлет-декомпозиция или база нечетких правил, чтобы компенсировать ее недостатки и повысить точность прогнозирования [72]. Все они повышают точность прогноза, но также требуют значительного вычислительного времени [73].

Некоторые более продвинутые, но сложные методы – деревья регрессии и их ансамбли, такие как метод «Случайного леса». В исследовательской работе [74] показано, что подход к регрессионному дереву решений работает лучше, чем классические инструменты регрессионного анализа и даже ANN, и уменьшает количество значимых факторов, влияющих на потребление энергии. Иногда сочетание преимуществ двух или более техник дает лучшую работу, чем один метод. Такой случай представлен формированием гибридной рекуррентной нейронной сети (RNN) с функциями вейвлет-активации, и показано улучшение точности прогноза по сравнению с простой нейронной сетью [75]. Другая предложенная гибридная модель прогнозирования использует генетический алгоритм для оптимизации выбора модели искусственного интеллекта (ИИ) и, как сообщается, превосходит несколько продвинутых моделей ИИ [76].

Из обзора литературы по интеллектуальным моделям прогнозирования можно сделать вывод, что будущий вклад заключается в разработке и тестировании новых подходов к искусственному интеллекту. Модели прогнозирования,

основанные на ANN, имеют определенные недостатки, такие как локальная минимальная точка, проблемы с подгонкой и т.д. [77]. Они могут быть преодолены с помощью передовых гибридных интеллектуальных моделей, таких как SVM, экстремальная обучающая машина (ELM) и адаптивная система нейро-нечеткого вывода (ANFIS), которые, как сообщается, демонстрируют хорошую производительность с точки зрения точности прогноза [78]. Однако вычислительные требования большинства из этих моделей довольно утомительны, особенно если задействовано обучение с помощью метода оптимизации [79]. Это недостаток для реальных приложений, таких как конкурентные энергетические рынки, где процесс торгов происходит очень быстро с большим количеством претендентов [80].

2.3 Влияние метеорологических факторов и изменения климата на выбор модели прогнозирования

В задачах прогнозирования генерации от различных возобновляемых источников энергии наиболее важным фактором является изменение климата, протекающее, как правило, с низкой скоростью, но которое в последние годы ускоряется. Кроме того, изменение климата обычно происходит постепенно, но могут возникнуть непредсказуемые краткосрочные изменения. В результате модели машинного обучения, построенные с использованием метеорологических факторов с хорошими результатами во время тестирования, могут значительно снизить точность прогнозирования через некоторое время из-за изменения внешних условий. Риск такого сценария особенно высок в отсутствие данных за долгую историю наблюдений. Поэтому необходимо, во-первых, проанализировать влияние на точность модели даже тех факторов, преимущества которых кажутся очевидными. Во-вторых, разумно применять модели, которые не подвержены переобучению и могут быть переобучены на относительно небольшом наборе данных при изменении климатических условий, адаптированных к изменениям. В результате в выборке, используемой для обучения, валидации и тестирования

модели, содержащий наблюдения за несколько лет, климатические условия могут существенно отличаться от условий всего лишь через один год. В результате какой бы сложной не была предиктивная модель и какие бы хорошие результаты она не показывала на используемой в исследовании выборке данных, качество ее работы через некоторое время может оказаться непригодным.

Поэтому в таких задачах необходимо использовать модели, которые могут быть быстро, автоматически дообучены или обучены с нуля на новых данных. Сложные модели, такие как глубокие нейронные сети, крупные ансамблевые модели часто требуют больших вычислительных затрат на процесс обучения, а из-за высокого влияния гиперпараметров на этот процесс необходимо участие человека-эксперта, который бы управлял процессом обучения, оценивал бы – нет ли эффекта переобучения и т.д. Можно исключить эксперта из данного процесса, используя различные методы автоматического подбора гиперпараметров [81-83], но это в разы увеличит требования к вычислительным ресурсам. Если ресурсов недостаточно, то многократное обучения модели с разными значениями гиперпараметров будет занимать недопустимо длительное время.

Таким образом, не всегда сложные по используемым алгоритмам и по своему размеру (например, общее число нейронов сети или глубина и число деревьев решений в ансамбле) модели могут быть применимы на практике в задачах прогнозирования для возобновляемой энергетики по следующим причинам:

1. Отсутствие необходимого объема ретроспективных данных для обучения. Причины:

- резкое изменение климатических, геологических, технологических факторов приводит к тому, что наблюдения за прошлые годы оказываются нерелевантными, так как меняются сами скрытые зависимости выходной прогнозируемой величины от наблюдаемых величин;

- может отсутствовать оцифрованная история наблюдений по всем требуемым факторам, особенно в развивающихся странах.

2. Высокие затраты на разработку, внедрение, эксплуатацию и непрерывное дообучение моделей. В рамках исследований ученые с большим опытом, высокой

квалификацией и доступом к вычислительным мощностям могут, вложив свое время, и навыки разработать модель с высокой точностью. Но для ее применения на практике требуется еще и создание автоматизированной информационной системы, закупка вычислительного оборудования, обслуживание как программной, так и аппаратной частей и наличие высококвалифицированных специалистов, которые способны быстро определить, что модель нуждается в модификации и выполнить нужное дообучение модели или ее обучения с нуля.

3. В модели могут быть использованы метеорологические данные, данные аэросъемки, спутникового мониторинга и т.д., которые могут быть в открытом доступе для ученых, но не для использования предприятием, производящем электроэнергию. Предприятию нужно будет платить за использование этих данных.

2.4 Сравнение методов прогнозирования электропотребления и генерации

Произведен анализ современных методов и моделей среднесрочного (от несколько часов до несколько недель вперед) прогнозирования электропотребления и генерации от ВИЭ, а именно статистических методов и методы на основе машинного обучения.

Для наглядности рассмотренных методов ниже в таблице 2.1 приводятся их достоинства и недостатки.

Таблица 2.1 – Достоинства и недостатки методов среднесрочного прогнозирования электропотребления и генерации

Обозначение и источник	Подход	Достоинства и недостатки
A1 [18 – 20]	На основе применения статистических методов	(+) Простота, гибкость и согласованность анализа проектирования, (-) Трудность определения нелинейных функциональных зависимостей
A2 [21;22]		(+) Простота и прозрачность моделирования, (-) Большое количество параметров модели, низкий уровень адаптивности, а также линейность
A3 [23-25]		(+) Проста вычисления и гибкость описания разных изменений, (-) Параметры сглаживания и начальные условия, не подходят для долгосрочного прогнозирования
A4 [26;27]		(+) Уникальность и простота способа определения решения в случае нормального распределения, (-) Трудность в вычислении оценок, связанных с решением уравнений правдоподобия
A1 - A4 [28-35]		Вывод: (+) Приемлемые результаты для коротких временных интервалов до 48 часов, (-) Прогнозирование становится очень неустойчивым для более длительных периодов времени (-) Для преобразования данных в стандартную форму требуется предварительная обработка данных временных рядов.
B1 [37-39]	На основе машинного обучения	(+) Позволяют учесть большое количество дополнительных факторов, влияющих на качество прогнозирования, (-) Медленное обучение и повышенные требования к программно-аппаратным ресурсам
B2 [40]		(+) Высокая скорость обучения и статистическая обоснованность процесса обучения, масштабируемость, (-) Неоднозначность алгоритма построения дерева, склонность деревьев к переобучению и невозможность интерпретации моделей
B3 [41]		(+) Простота реализации и прозрачность процедуры работы. (-) Необходимость выбора гипер-параметров, в первую очередь числа соседей k , а также метрики расстояний
B4 [42;50]		(+) Позволяет строить сложные модели для выявления сложных зависимостей за счет масштабирования: увеличения числа слоев модели, применения различных элементов обработки данных (-) Локальная минимальная точка, проблемы с подгонкой
B1 – B4 [43]		Вывод: (+) Обладает нелинейностью, то есть позволяет устанавливать нелинейную зависимость между будущими и прошлыми значениями, адаптивность и масштабируемость, (-) Отсутствует прозрачность моделирования, сложность выбора архитектуры сети, предъявляет высокие требования к объему и качеству обучающей выборки

Выводы по главе 2

В настоящее время в энергосистемах используется большое количество разновидностей методов среднесрочного прогнозирования электропотребления и генерации, начиная от простых статистических методов, и заканчивая искусственными нейронными сетями.

Согласно проведенному анализу существующих методов для задач среднесрочного прогнозирования электропотребления и генерации целесообразно использовать методы, основанные на машинном обучении.

ГЛАВА 3 РАЗРАБОТКА МОДЕЛИ СРЕДНЕСРОЧНОГО ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ И ГЕНЕРАЦИИ С УЧЕТОМ МЕТЕОРОЛОГИЧЕСКИХ ФАКТОРОВ

3.1 Среднесрочная модель прогнозирования энергопотребления

Среднесрочное прогнозирование чрезвычайно важно для поставщиков энергии и других участников рынков производства, передачи, распределения и электроэнергии. Это помогает принимать решения, в том числе о покупке и производстве электроэнергии в электроэнергетической системе. Основная роль прогнозирования электропотребления в такой системе заключается в повышении точности планирования производства и купли-продажи необходимого объема электроэнергии [84-86].

Одним из возможных подходов к повышению точности прогнозов является использование дополнительных этапов фильтрации и предобработки данных, в том числе с помощью кластеризации. Кластеризация позволяет строить вместо единой модели для всех условий функционирования ряд отдельных моделей, каждая из которых лучше адаптирована к определенным метеорологическим условиям. При этом для текущих условий среди всех ретроспективных данных формируется выборка из наиболее похожих записей, затем на этих данных обучается модель. Чаще всего кластеризация метеорологических условий в задачах электроэнергетики связана с прогнозированием генерации возобновляемых источников энергии.

Кластеризация может быть использована и для фильтрации данных. В [87] авторы описывают применение кластеризации метеорологических параметров для обнаружения выбросов. При анализе электропотребления также применяется кластеризация. Например, для разделения потребителей по профилям электропотребления [88,89].

В данном разделе работы кластеризация применена для разделения выборки ретроспективных данных на несколько категорий, различающихся по

метеорологическим условиям. Для каждого кластера создается своя модель прогнозирования нагрузки. Также отличием предложенного подхода является алгоритм выбора оптимального числа предыдущих дней, метеорологические данные которых используются для кластеризации. Для этого применен алгоритм снижения размерности массива данных Principal Component Analysis (Анализ Основных Компонентов) (РСА), который позволяет визуально оценить, как распределились кластеры в пространстве признаков. Также алгоритм РСА позволил упростить прогнозные модели за счет агрегирования метеорологических данных.

3.1.1 Описание выборки данных

Поскольку дефицит электроэнергии в ГБАО, вырабатываемой ГЭС, наблюдается только в холодный период (с ноября по март включительно), рассмотрены данные только за эти месяцы (Рисунок 3.1 и 3.2).

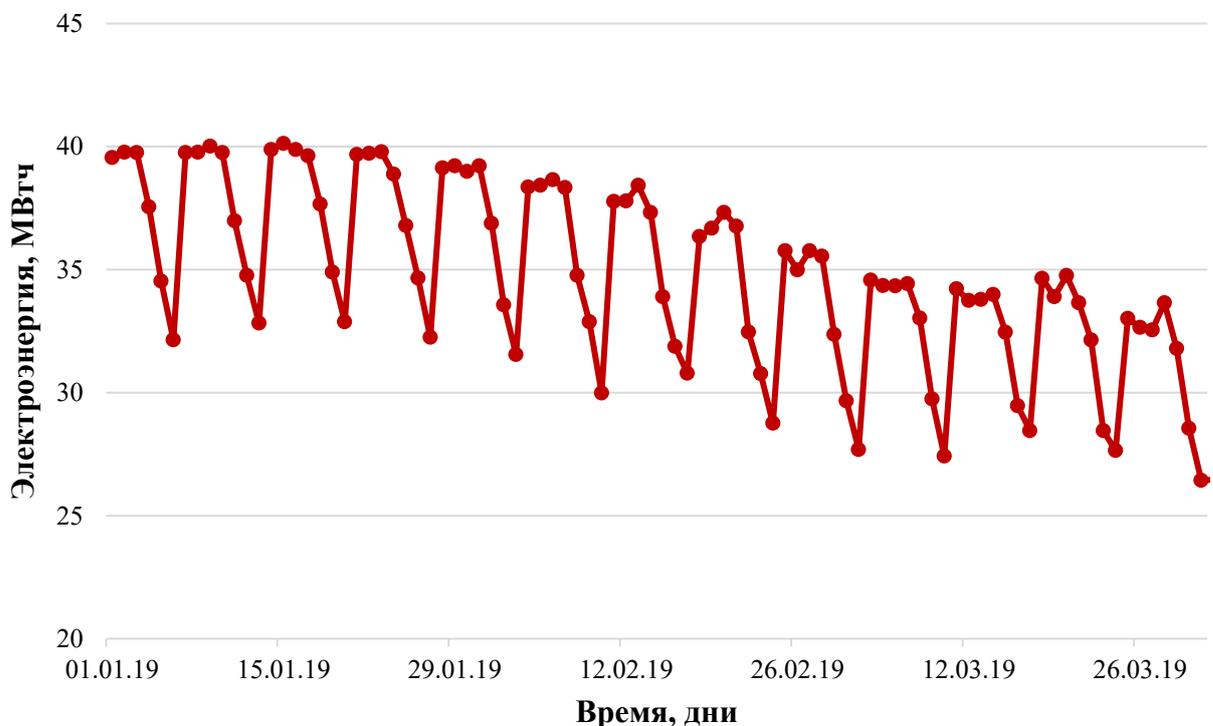


Рисунок 3.1 – Графики электропотребления за январь-март 2019 г.

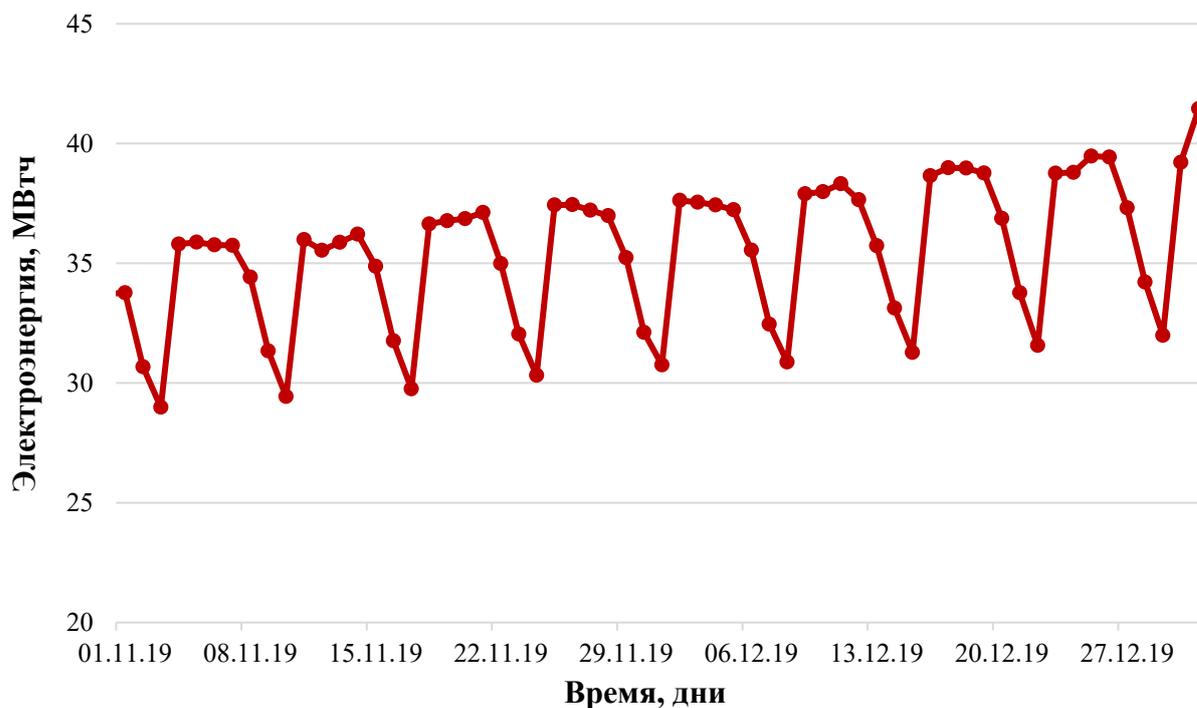


Рисунок 3.2 – Графики электропотребления за ноябрь-декабрь 2019 г.

Основная гипотеза данного исследования заключается в возможности повысить точность среднесрочного прогнозирования электропотребления за счет использования метеорологических данных и кластеризации метеорологических условий.

Кроме того, в рассматриваемой энергосистеме нагрузка меняется в выходные дни, поэтому в выборку добавлен номер дня недели. Номера года, месяца и дня также рассмотрены в качестве признаков. Ниже приведен полный список признаков в исходной выборке данных:

- год;
- месяц;
- день;
- день недели;
- температура воздуха;
- относительная влажность воздуха;
- скорость ветра;
- облачность.

Электропотребление (среднесуточная мощность нагрузки) является целевой переменной.

Выборка содержит записи с шагом в 1 сутки за 6 лет (2015–2020 гг.). Таким образом, выборка содержит 908 строк и 9 столбцов.

Коэффициенты корреляции между признаками и электропотреблением приведены в Таблице 3.1. Видно, что наибольшая зависимость наблюдается между электропотреблением и годом и между электропотреблением и днем недели, но метеорологические факторы, такие как температура, влажность и скорость ветра, тоже оказывают влияние на электропотребление.

Таблица 3.1 – Коэффициент корреляции между признаками и электропотреблением

Признак	Коэффициент корреляции
Год	0.70
Месяц	-0.09
День	0.01
День недели	-0.51
Температура	-0.33
Влажность	0.20
Скорость ветра	-0.37
Облачность	0.01

3.1.2 Постановка задачи прогнозирования

Необходимо построить модель среднесрочного (на неделю вперед) прогнозирования электропотребления энергосистемы:

$$P_i^* = f(W_{i-7}, W_{i-8}, \dots, W_{i-6-d}) \quad (3.1)$$

где

P_i^* – прогнозируемое электропотребление в i -е сутки;

f – модель;

d – число предыдущих дней, использованных для прогнозирования

W_j – вектор значений используемых признаков в j -е сутки, включающий в себя и P_j .

В данном разделе работы не используются прогнозы метеорологических факторов, поскольку в этом случае система окажется зависимой от точности среднесрочных прогнозов метеопровайдеров. Для прогнозирования использованы данные за s предыдущие сутки (с учетом того, что горизонт планирования составляет неделю).

Точность прогноза оценивается на тестовой выборке с использованием метрик mean absolute error (MAE), mean absolute percentage error (MAPE) и root mean squared error (RMSE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - P_i^*| \quad (3.2)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{P_i - P_i^*}{P_i} \right| \quad (3.3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - P_i^*)^2} \quad (3.4)$$

где

P – истинное значение электропотребления;

P^* – прогнозное значение электропотребления;

n – размер тестовой выборки в сутках.

3.1.3 Кластеризация метеорологических условий

При построении моделей машинного обучения необходимо сделать выбор: создавать единую модель, которая будет решать всю задачу целиком, или разбить задачу на этапы и создавать отдельные модели для каждой из них.

Создание единой модели позволяет выполнять ее обучение для достижения конечной цели, в то время как при разбиении задачи на этапы каждая модель строится для достижения подцели. Затем результаты объединения таких моделей

могут оказаться не столь качественными, как при обучении единой модели, сфокусированной на конечном результате. С другой стороны, в задачах планирования и управления в электроэнергетике очень важна интерпретируемость моделей машинного обучения, гарантия их робастности и снижение риска неожиданного поведения [90].

Кроме того, нужно учитывать ограниченность выборки данных, чем меньше выборка, тем менее надежными являются выводы, полученные с использованием сложных моделей. Разбиение задачи на части позволяет применять более простые модели на каждом этапе и контролировать их. Ограничение размера модели может сделать ее более устойчивой к переобучению, более робастной и повысить ее обобщающую способность [91].

Поэтому в данном разделе работы использована следующая архитектура системы прогнозирования на базе машинного обучения:

- 1) уровень предобработки данных;
- 2) уровень кластеризации данных;
- 3) уровень прогнозирования, состоящий из m независимых нейронных сетей, по числу кластеров.

3.1.4 Прогнозирование электропотребления

Суть предложенного подхода заключается в разбиении метеорологических условий на m кластеров с помощью машинного обучения без учителя (алгоритм k -means [92]). Для каждого кластера выполняется построение своей нейросетевой модели. При работе системы входные данные относятся к одному из кластеров и затем для получения прогноза применяется та модель, которая была построена для этого кластера.

Такой подход имеет следующие преимущества:

- этап кластеризации выполняется без учителя, значит, по определению исключен риск подгонки под нужный результат;
- позволяет визуально проверить логичность результата кластеризации,

применив после кластеризации метод главных компонент и визуализировав полученные кластеры на плоскости;

- построение разных моделей для отдельных метеорологических позволяет повысить точность прогноза, поскольку каждая модель оказывается сфокусированной на определенных условиях работы;

- отдельные модели могут быть более компактными, чем единая – следовательно, для их обучения и тестирования можно использовать меньше данных, а их работа будет более предсказуемой.

Исходная выборка данных может быть записана как

$$\{(W_i, P_i)\}, i = 1 \dots n$$

На этапе предобработки она переводится в следующий вид, соответствующий выражению (1):

$$\{(W_{i-7}, W_{i-8}, \dots, W_{i-6-d}, P_i)\}$$

Число d определяет количество предыдущих суток, которые будет использоваться для построения прогноза. При d с равном 1, полученный датасет будет иметь 10 столбцов (9 признаков и целевую переменную), при d равном 2 – 19 и т.д.

Кроме того, на этапе предобработки данных выполняется Min-Max нормализация, чтобы все значения признаков оказались в диапазоне от 0 до 1.

Для кластеризации использован алгоритм k-средних, так как он является general-purpose алгоритмом. Чтобы проверить корректность кластеризации и выбрать наиболее подходящее число кластеров, в работе использован Principal Component Analysis (PCA) [93]. Кластеризация выполняется по признакам $(W_{i-7}, W_{i-8}, \dots, W_{i-6-k})$. Затем к данным применяется PCA с переводом размерности признаков в двумерную плоскость. Точки на плоскости затем отмечаются цветами в соответствии с метками кластеров. Четкость границ между полученными кластерами показывает, насколько кластеризация выполнена корректно.

Рисунки 3.3 и 3.4 показывают полученные результаты. Изначально мы использовали 4 кластера, это количество выбрано исходя из визуального анализа результатов PCA. Но затем на этапе построения регрессионных моделей, было

обнаружено, что три кластера дают лучший результат. Видно, что при числе используемых дней менее пяти границы между кластерами являются менее четкими. Использование более семи дней также нецелесообразно, так как может привести к переобучению регрессионных моделей, поскольку признаков будет слишком много относительно числа дней в выборке данных.

Таким образом, на следующих этапах использованы посуточные данные за неделю и 3 или 4 кластера.

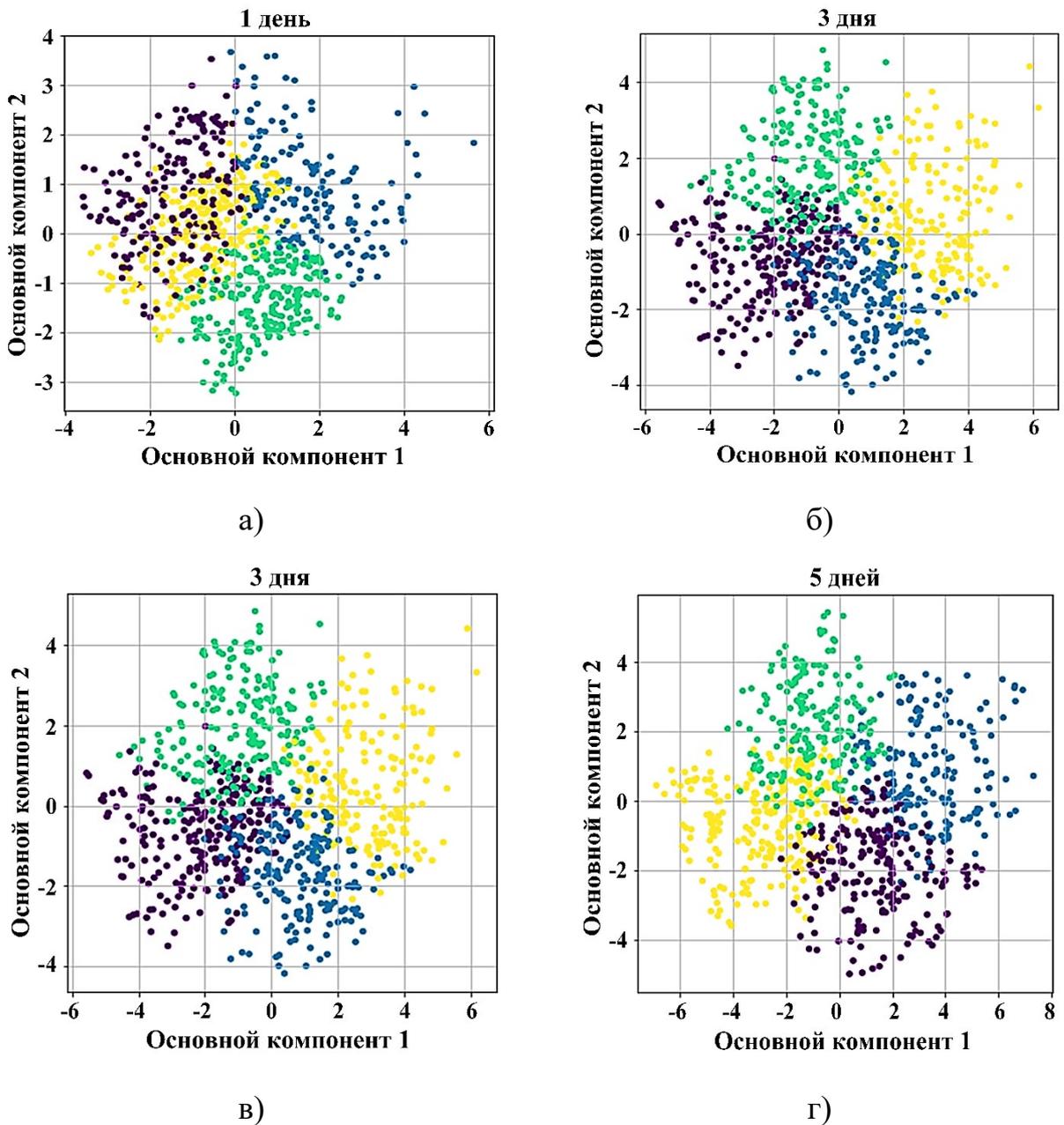


Рисунок 3.3 – Визуализация 4-х кластеров (а – данные за три дня, б – за четыре, в – за пять, г – за семь), цвет означает кластер

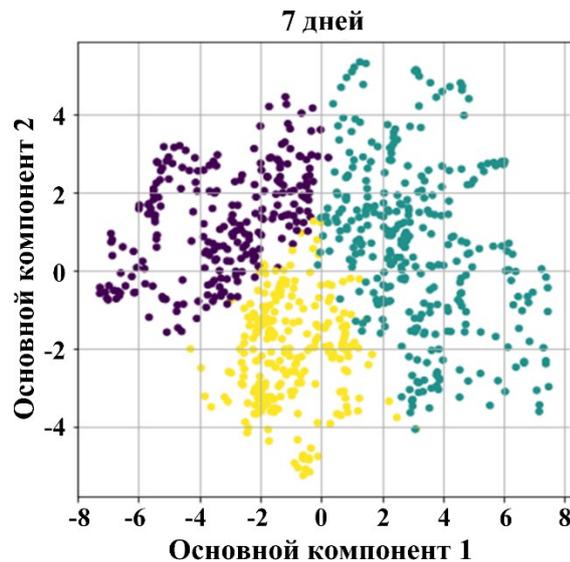


Рисунок 3.4 – Визуализация 3-х кластеров при использовании данных за семь дней

3.1.5 Нейросетевая модель и результаты

Поскольку используется множество разнородных факторов, в качестве базового метода построения регрессионной модели выбрана искусственная нейронная сеть (ANN). Выборка данных не настолько большая, чтобы использовать более сложные архитектуры, такие как сверточные нейронные сети или рекуррентные. Поэтому использован многослойный перцептрон следующего вида:

- 1) входной слой;
- 2) скрытый слой из 32 нейронов с функцией активации ReLU;
- 3) скрытый слой из 16 нейронов с функцией активации ReLU;
- 4) скрытый слой из 8 нейронов с функцией активации ReLU;
- 5) выходной нейрон с сигмоидальной функцией активации.

Обучение выполняется с помощью алгоритма Adam [94]. Модель реализована с помощью Python библиотек Keras и Tensorflow.

Проведено сравнение предложенного подхода с единой нейросетевой моделью (без учета кластеров), с авторегрессионной моделью (AR), не использующей метеорологические данные и с линейной регрессией (LR),

использующей метеорологические данные. Результаты показаны в Таблице 3.2. В ней представлены результаты на тестовой части выборки (10 % более новых данных, а 90 % более старых использованы для обучения).

Таблица 3.2 – Результаты прогнозных моделей с использованием метеорологических данных

Модель	Число кластеров	Номер кластера	MAE	MAPE	RMSE
			МВт	%	МВт
ANN	-	-	1.023	7.48	1.27
ANN	4	1	0.653	4.73	0.81
ANN	4	2	0.779	8.16	0.97
ANN	4	3	1.495	9.60	1.68
ANN	4	4	0.905	8.81	1.34
ANN	4	средне взвешенное	0.939	7.70	1.19
ANN		1	1.067	7.15	1.40
ANN	3	2	0.734	6.13	0.92
ANN	3	3	0.776	6.90	1.00
ANN	3	средне взвешенное	0.846	6.67	1.09
LR	-	-	0.774	5.82	0.99
LR	3	1	0.604	5.03	1.04
LR	3	2	0.751	4.86	0.73
LR	3	3	0.397	3.52	0.56
LR	3	средне взвешенное	0.586	4.53	0.77
AR	-	-	0.803	6.11	0.96

Из полученных результатов сравнения моделей можно сделать следующие заключения:

- использование кластеризации по метеорологическим условиям повышает точность как для линейной регрессии, так и для искусственных нейронных сетей; снижение ошибки составило для ANN с 1.023 МВт до 0.846 МВт (17 %), для LR с 0.774 до 0.586 (24 %);

- но в то же время может наблюдаться обратный эффект, если число кластеров больше оптимального, поскольку в каждом из кластеров окажется недостаточно данных для обучения модели;

- использование метеорологических условий снижает ошибку прогноза, разница ошибки между авторегрессионной моделью, не использующей метеорологические параметры и лучшей моделью, использующих их, составила 0.217 МВт (27 %);

- несколько неожиданным оказалось существенное преимущество линейной регрессии над нейросетевыми моделям.

Последний пункт можно объяснить двумя факторами:

1) использование признаков, полученных с помощью PCA, в качестве дополнительных входных признаков позволило во много выполнить агрегирование метеорологических факторов еще до применения регрессионных моделей;

2) объем данных недостаточно велик для эффективного применения нейронных сетей, для их более эффективной работы, вероятно, требуются данные за период 10-20 лет.

Необходимо подчеркнуть, что полученный результат позволяет создать надежную и очень простую модель на базе машинного обучения, с минимальным использованием сложных и слабо контролируемых методов машинного обучения.

Итоговая модель выглядит следующим образом.

1. На этапе построения прогнозной системы выполняется:

- 1) кластеризация алгоритмом k-means на три кластера;
- 2) применение PCA для формирования двух наиболее значимых признаков, агрегирующих метеорологические параметры;
- 3) построение для каждого кластера своей линейной регрессионной модели.

2. На этапе эксплуатации выполняется:

- 1) определение, как какому кластеру относятся текущие условия по данным метеорологических наблюдений в течение предыдущей недели;
- 2) применение PCA для получения агрегированных признаков;

3) применение линейной регрессионной модели для получения прогноза потребления.

На рисунке 3.5 показано сопоставление истинного графика нагрузки и прогнозного для фрагмента тестовой выборки.

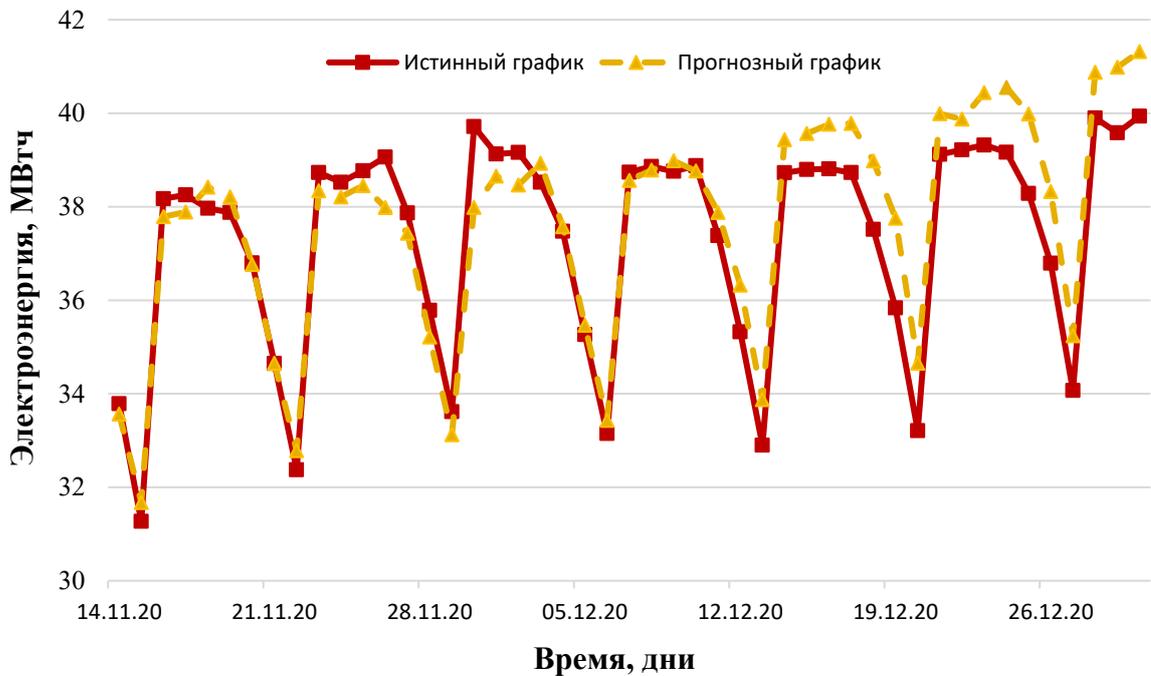


Рисунок 3.5 – Сравнение истинного и прогнозного графиков электропотребления

3.2 Среднесрочная модель прогнозирования притока воды для планирования генерации

В последние десятилетия малые ГЭС (МГЭС) получили значительное развитие из-за их малых масштабов, низких инвестиций и их быстрой отдачи, экологически безопасных энергетических технологий и содействия местному экономическому развитию [95-98]. При объединении таких МГЭС в изолированные энергосистемы в условиях дефицита генерирующих мощностей возникает потребность в оптимизации состава генерирующего оборудования [99,100]. Ввиду существенной зависимости генерации от неуправляемого притока воды решение задачи по определению ее запаса в среднесрочной перспективе становится первостепенной. Решение такой задачи позволит координировать

режимы работы МГЭС, обеспечивая балансирование между необходимой генерацией и достаточными ресурсами.

Возможность сезонного регулирования определяется возможностью накопления значительных объемов воды, достаточных для регулирования генерации в течение долгосрочного периода. В настоящее время прогнозирование генерации ГЭС выполняется на основе определения перспективного притока в водохранилище и стока реки [101,102]. В случае применения МГЭС возможность прогнозирования снижается, возникают дополнительные препятствия для устойчивого определения прогнозных показателей. Такие станции, как правило, размещаются на стоках рек с малой регулирующей мощностью, что не позволяет обеспечить долгосрочное устойчивое регулирование. Высокая неопределенность, большое пространственно-временное разнообразие, сильная зависимость от внешних условий затрудняют применение стандартных моделей прогнозирования, используемых для ГЭС большой мощности [103,104]. Возможность генерации МГЭС определяется такими внешними факторами, как гидрологические и климатические условия, сезонность, высота установки генерирующего оборудования над уровнем моря. Поэтому прогнозирование генерации на МГЭС является достаточно трудоемкой нелинейной и многокритериальной задачей.

3.2.1 Постановка задачи и исходные данные

В данном разделе работы использованы данные о притоке воды озера Яшилькуль за 2010–2013 гг. (Рисунок 3.6) и 2015–2019 гг. (Рисунок 3.7), 2014 год отсутствует в истории наблюдений. Каждый год представлен только за период с ноября по март (151 или 152 дня), период нехватки электроэнергии из-за вышеуказанных проблем.

Из сравнения Рисунок 3.6 и Рисунок 3.7 можно заметить, насколько сильно меняется сам характер прогнозируемой величины в разные периоды. Например, ни в один из годов 2010–2013 не наблюдалось резкого повышения в декабре, что произошло в 2015 году.

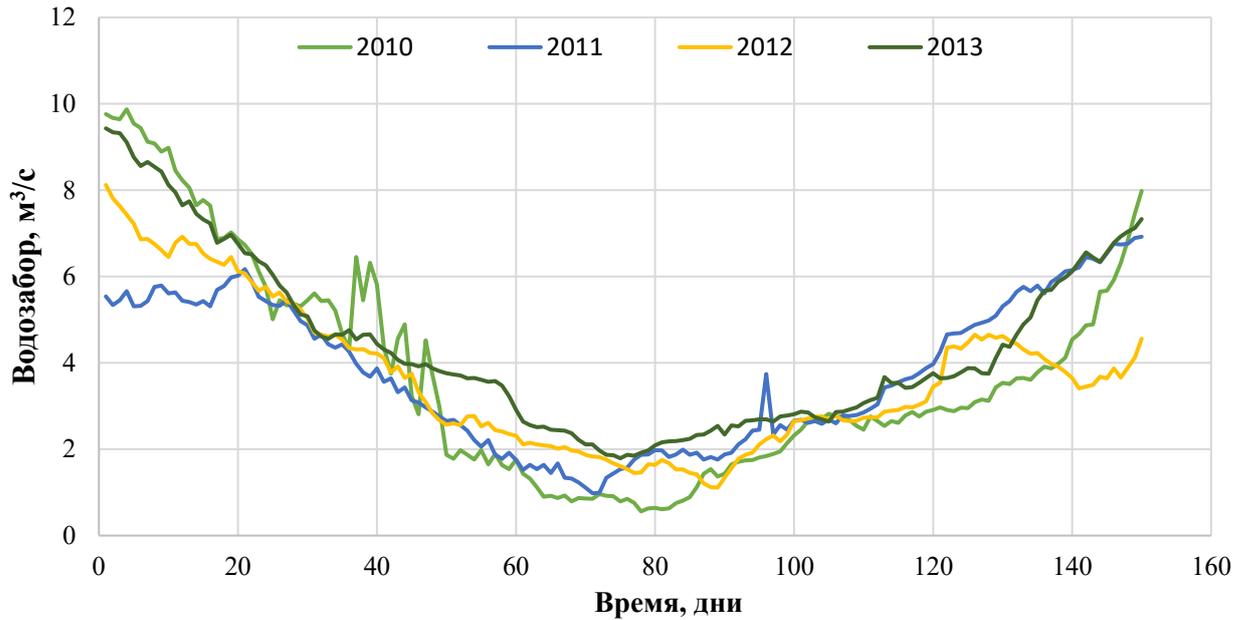


Рисунок 3.6 – Графики притока воды 2010-2013 гг.

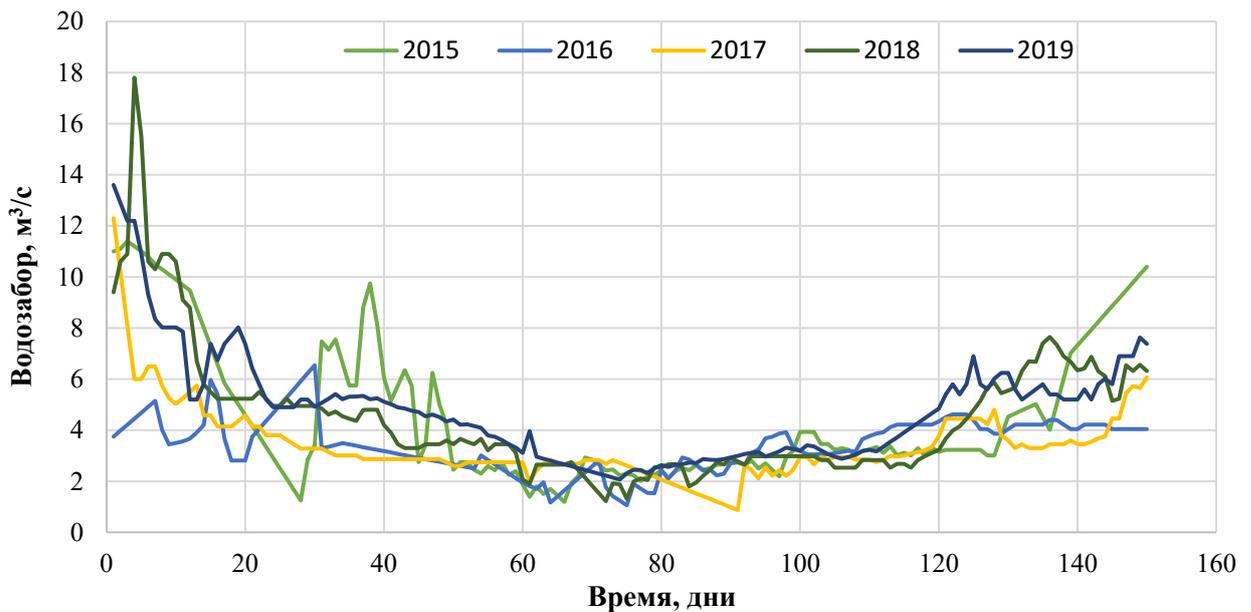


Рисунок 3.7 – Графики притока воды 2015-2019 гг.

В данном разделе работы рассмотрен подход по использованию компактных моделей машинного обучения для среднесрочного прогнозирования притока воды, отличием которого является полная автономность и автоматизация обучения, исключение риска переобучения и отсутствие необходимости в каких-либо дополнительных данных кроме самого временного ряда.

Рассмотрена задача прогнозирования притока воды за сутки с горизонтом планирования в одну неделю. Задача прогнозирования в данной работе сформулирована следующим образом. Построение модели, выполняющей преобразование входных данных (признаков, описанных ниже) X в прогноз суточного притока на неделю вперед Y^* :

$$Y^* = f(X) \quad (3.5)$$

Показателем точности прогноза в данном разделе работы выбран корень из среднего квадрата ошибки (RMSE), который фокусируется на больших ошибках:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^* - y_i)^2} \quad (3.6)$$

где n – число дней в выборке; y_i^* – прогнозное значение в i -е сутки; y_i – истинное значение в i -е сутки.

Для наглядности и отсутствия привязки к конкретному значению притока данного объекта используется ошибка в процентах, $nRMSE$, то есть $RMSE$ деленный на среднее значение притока.

$$nRMSE = \frac{RMSE}{Y} 100\% \quad (3.7)$$

Альтернативный показатель MAPE, дающий ошибку в процентах, в данной задаче плохо подходит, поскольку он очень чувствителен к малым по абсолютной величине ошибками, которые могут оказываться очень высокими в процентах из-за деления ошибки на малую величину истинного значения.

3.2.2 Используемый метод и результаты

Формирование выборки данных для применения машинного обучения

Из исходных данных была сформирована выборка, структура которой и примеры значений данных показана в Таблице 3.3. В ней w – суточные данные притока воды. Таким образом, для прогноза на некоторые сутки с номером j на неделю вперед используются данные двух недель (от суток $j-21$ до суток $j-8$). Дни

и месяцы отсчитывались от начала периода (ноябрь – первый месяц), номер года – от начала наблюдений, 2010.

На Рисунке 3.8 показана матрица корреляций признаков по критерию Спирмена. Видно, что корреляция прогнозируемого притока суток (Y) и даты (годы, месяц, день) слабая, более сильная корреляция наблюдается с притока в предыдущие недели. Видно, что приток за сутки, которые отстают от прогнозной на 2-3 недели (w_1-w_5) слабо коррелируют с прогнозируемой величиной, поэтому они были исключены из выборки для дальнейшей работы.

Таблица 3.3 – Структура выборки, используемой при построении моделей

День	Месяц	Год	w_1	w_2	...	w_{14}	$Y(w_{22})$
22	1	1	9.76	9.67	...	7.65	6.52
23	1	1	9.67	9.64	...	7.7	6.12
...
151	5	10	5.00	5.80	...	5.80	7.86

3.2.3 Описание экспериментов

Целью вычислительных экспериментов было определить, насколько модель, обученная и проверенная на данных одного периода, может быть использована в последующие несколько лет и проверить, можно ли повысить точность прогноза, если вместо обучения более сложной модели использовать модели машинного обучения достаточно компактные, чтобы риск их переобучения и подгонки был минимален, а скорость обучения достаточно быстрая, чтобы провести исследования, имитирующее ежедневное обучение модели на новых данных.

Для достижения цели были проведены следующие вычислительные эксперименты.

1. Построение и анализ многослойной рекуррентной нейронной сети долговременной памяти (пример сложной модели) (LSTM).

1.1. Обучение, настройка гиперпараметров и валидация модели на данных 2010–2013 гг., тестирование на данных 2014 г.

1.2. Обучение и валидация модели на данных 2010–2014 гг., тестирование по-

отдельности на данных 2015–2019 гг. Гиперпараметры модели использованы такие же, которые найдены на шаге 1.1.

2. Построение и анализ моделей с компактным размером, что ограничивает их переобучение. Используются классические многослойный перцептрон [105], адаптивный бустинг (AdaBoost) [106], случайный лес (Random Forest) [107].

Шаги 2.1 и 2.2 такие же, как 1.1 и 1.2.

2.3. Самоадаптация. Имитация режима постоянного повторного обучение (retraining) моделей. Чтобы получить прогноз на дату D_1 , модель обучается на всех данных с 2010 года до даты $D_0 = D_1 - 1$ неделя. И так делается для каждого суток года. При эксплуатации это соответствует режиму, при котором модель обучалась бы каждый день.

Шаг 2.3 для рекуррентной многослойной модели не выполняется, поскольку ее обучение длится на два порядка дольше, так что обучать ее для каждого дня слишком долго.

год	1.00	-0.00	-0.00	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.15	0.15	0.15	0.15	0.15	0.15	0.16
день	-0.00	1.00	0.97	-0.46	-0.44	-0.41	-0.38	-0.36	-0.33	-0.30	-0.27	-0.23	-0.20	-0.17	-0.14	-0.11	-0.07	0.20	
месяц	-0.00	0.97	1.00	-0.47	-0.44	-0.42	-0.39	-0.37	-0.33	-0.30	-0.27	-0.24	-0.21	-0.17	-0.14	-0.11	-0.07	0.22	
w1	0.14	-0.46	-0.47	1.00	0.98	0.95	0.93	0.91	0.89	0.86	0.84	0.82	0.80	0.77	0.74	0.72	0.69	0.45	
w2	0.14	-0.44	-0.44	0.98	1.00	0.98	0.95	0.93	0.91	0.89	0.86	0.84	0.82	0.80	0.77	0.74	0.71	0.47	
w3	0.14	-0.41	-0.42	0.95	0.98	1.00	0.98	0.95	0.93	0.91	0.89	0.86	0.84	0.82	0.79	0.77	0.74	0.50	
w4	0.14	-0.38	-0.39	0.93	0.95	0.98	1.00	0.98	0.95	0.93	0.91	0.89	0.86	0.84	0.82	0.79	0.76	0.53	
w5	0.14	-0.36	-0.37	0.91	0.93	0.95	0.98	1.00	0.98	0.95	0.93	0.91	0.88	0.86	0.84	0.81	0.79	0.56	
w6	0.14	-0.33	-0.33	0.89	0.91	0.93	0.95	0.98	1.00	0.98	0.95	0.93	0.90	0.88	0.86	0.83	0.81	0.58	
w7	0.14	-0.30	-0.30	0.86	0.89	0.91	0.93	0.95	0.98	1.00	0.97	0.95	0.93	0.90	0.88	0.85	0.83	0.61	
w8	0.15	-0.27	-0.27	0.84	0.86	0.89	0.91	0.93	0.95	0.97	1.00	0.97	0.95	0.92	0.90	0.88	0.85	0.64	
w9	0.15	-0.23	-0.24	0.82	0.84	0.86	0.89	0.91	0.93	0.95	0.97	1.00	0.97	0.95	0.92	0.90	0.88	0.67	
w10	0.15	-0.20	-0.21	0.80	0.82	0.84	0.86	0.88	0.90	0.93	0.95	0.97	1.00	0.97	0.95	0.92	0.90	0.70	
w11	0.15	-0.17	-0.17	0.77	0.80	0.82	0.84	0.86	0.88	0.90	0.92	0.95	0.97	1.00	0.97	0.95	0.92	0.72	
w12	0.15	-0.14	-0.14	0.74	0.77	0.79	0.82	0.84	0.86	0.88	0.90	0.92	0.95	0.97	1.00	0.97	0.95	0.75	
w13	0.15	-0.11	-0.11	0.72	0.74	0.77	0.79	0.81	0.83	0.85	0.88	0.90	0.92	0.95	0.97	1.00	0.97	0.78	
w14	0.15	-0.07	-0.07	0.69	0.71	0.74	0.76	0.79	0.81	0.83	0.85	0.88	0.90	0.92	0.95	0.97	1.00	0.80	
Y	0.16	0.20	0.22	0.45	0.47	0.50	0.53	0.56	0.58	0.61	0.64	0.67	0.70	0.72	0.75	0.78	0.80	1.00	
	год	день	месяц	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	w11	w12	w13	w14	Y	

Рисунок 3.8 – Матрица коэффициентов корреляции

3.2.4 Модели на базе машинного обучения и результаты

Спецификация моделей приведена в Таблице 3.4. Реализации моделей MLP, AdaBoost и Random Forest взяты из библиотеки машинного обучения scikit-learning [108], mLSTM создана с использованием блоков библиотеки Keras [109].

Результаты mLSTM модели приведены в Таблице 3.5, результаты MLP, AdaBoost и Random Forest в Таблицах 3.6–3.8, соответственно. В Таблицах 3.6–3.8 для шага 2.3 приведены показатели точности, усредненные по всем дня соответствующего года. Используемое обозначение $2010-D_0(2015)$ означает, что при обучении модели использовались все данные с начала периода до той даты в 2015 год. То есть использован не один фиксированный период, а для каждого дня в 2015 году сформирована своя выборка от ноября 2010 года до суток за неделю до дня, на который дается прогноз.

Таблица 3.9 показывает результаты в более удобном формате для сравнения моделей и режимов их применения, те же результаты представлены на Рисунке 3.10. Рисунок 3.11 показывает результаты AdaBoost с онлайн обучением и без него.

Таблица 3.4 – Используемые модели

Модель	Архитектура и гипер-параметры после настройки на данных 2010-2013 гг.
LSTM	См. Рис. 4; метод обучения Adam, размер пакета – 32
MLP	число скрытых слоев: 3, число нейронов в них 128, 64, 32; функция активации ReLU, метод обучения Adam
AdaBoost[Decision Trees] (AB)	Число деревьев 200, максимальная глубина 3
Random Forest (RF)	Число деревьев 100, максимальная глубина 4

Таблица 3.5 – Результаты LSTM, средняя ошибка в период 2015–2019 гг. 27.4 % или 4.93 м3/с

Период для обучения и валидации	Период для тестирования	Результаты на обучающей выборке, nRMSE, %	Результаты на тестовой выборке, nRMSE, %
2010–2012	2013	16.23	16.88
2010–2013	2015	15.07	37.24
	2016		26.90
	2017		24.81
	2018		25.65
	2019		22.25

Таблица 3.6 – Результаты MLP, средняя ошибка в период 2015–2019 гг. 30.35 % или 5.45 м3/с без обучения онлайн и 25.32 % или 4.56 м3/с при онлайн обучении

Период для обучения и валидации	Период для тестирования	Результаты на обучающей выборке, nRMSE, %	Результаты на тестовой выборке, nRMSE, %
2010–2012	2013	20.68	15.50
2010–2013	2015	18.22	52.54
	2016		28.25
	2017		29.30
	2018		22.78
	2019		18.89
2010– $D_0(2015)$	2015	24.53	40.71
2010– $D_0(2016)$	2016	26.30	29.88
2010– $D_0(2017)$	2017	26.48	22.60
2010– $D_0(2018)$	2018	25.57	22.25
2010– $D_0(2019)$	2019	24.95	11.14

Таблица 3.7 – Результаты AdaBoost[Decision Trees], средняя ошибка в период 2015–2019 гг. 24.43 % или 4.40 м3/с без обучения онлайн и 22.08 % или 3.97 м3/с при онлайн обучении

Период для обучения и валидации	Период для тестирования	Результаты на обучающей выборке, nRMSE, %	Результаты на тестовой выборке, nRMSE, %
2010–2012	2013	12.91	15.10
2010–2013	2015	12.70	46.02
	2016		26.05
	2017		21.17
	2018		17.77
	2019		13.88
2010 – $D_0(2015)$	2015	15.33	29.43

Продолжение таблицы 3.7

2010 – $D_0(2016)$	2016	18.80	25.29
2010 – $D_0(2017)$	2017	19.82	22.92
2010 – $D_0(2018)$	2018	20.31	20.16
2010 – $D_0(2019)$	2019	20.64	12.62

Таблица 3.8 – Результаты Random Forest, средняя ошибка в период 2015–2019 гг. 25.29 % или 4.55 м3/с без обучения онлайн и 22.63 % или 4.07 м3/с при онлайн обучении

Период для обучения и валидации	Период для тестирования	Результаты на обучающей выборке, nRMSE, %	Результаты на тестовой выборке, nRMSE, %
2010–2012	2013	11.01	13.20
2010–2013	2015	11.59	47.74
	2016		26.04
	2017		19.98
	2018		18.09
	2019		14.57
2010 – $D_0(2015)$	2015	16.51	35.22
2010 – $D_0(2016)$	2016	17.44	23.93
2010 – $D_0(2017)$	2017	17.83	19.41
2010 – $D_0(2018)$	2018	18.18	21.19
2010 – $D_0(2019)$	2019	18.34	13.40

Таблица 3.9 – Сопоставление результатов

Модель	Онлайн обучение	nRMSE, 2015	nRMSE, 2016	nRMSE, 2017	nRMSE, 2018	nRMSE, 2019	nRMSE, 2015-2019
mLSTM	нет	37.24	26.90	24.81	25.65	22.25	27.4
MLP	нет	52.54	28.25	29.30	22.78	18.89	30.35
MLP	да	40.71	29.88	22.60	22.25	11.14	25.32
AB	нет	46.02	26.05	21.17	17.77	13.88	24.43
AB	да	29.43	25.29	22.92	20.16	12.62	22.08
RF	нет	47.74	26.04	19.98	18.09	14.57	25.29
RF	да	35.22	23.93	19.41	21.19	13.40	22.63

Layer (type)	Output Shape	Params #
input (InputLayer)	[(None, 12, 1)]	0
lstm_1 (LSTM)	(None, 12, 64)	16896
dropout (Dropout 0.4)	(None, 12, 64)	0
lstm_2 (LSTM)	(None, 32)	12416
dense_1 (Dense)	(None, 32)	1056
activation_1 (ReLU)	(None, 32)	0
dense_2 (Dense)	(None, 16)	528
activation_2 (ReLU)	(None, 16)	0
dense_3 (Dense)	(None, 1)	17
activation_3 (Sigmoid)	(None, 1)	0
Total trainable params: 30,913		

Рисунок 3.9 – Архитектура LSTM модели

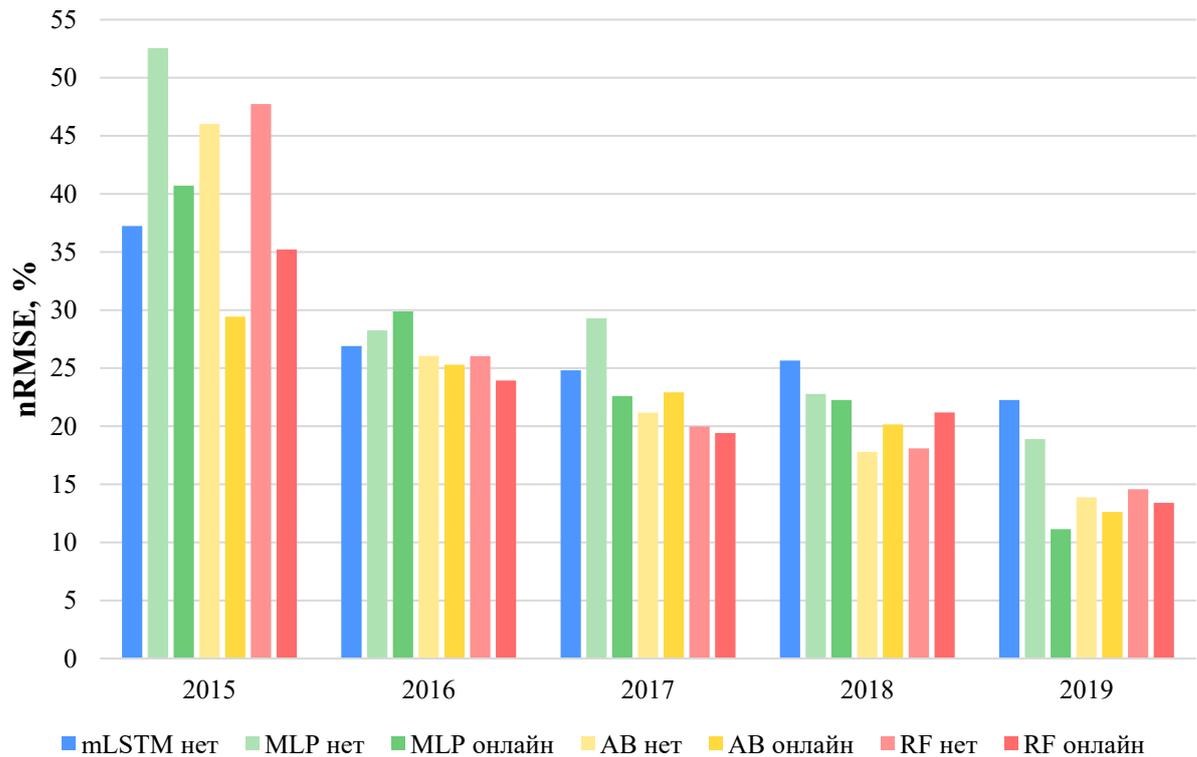


Рисунок 3.10 – Сопоставление результатов прогнозирования

3.2.5 Анализ полученных результатов

Результаты моделирования показывают следующее.

1. Из-за возможных климатических изменений в отдельный год график генерации может оказаться сильно отличающимся от нескольких предыдущих лет, поэтому приемлемая точность на тестовой выборке и отсутствие отклонения точности на тестовой и обучающей выборках еще не гарантируют работоспособности модели в следующие годы. Это видно из отличия ошибок всех моделей в 2013 г. от ошибок в 2015 г.

2. Если бы в истории наблюдений не было 2015 г., а точнее его ноября и декабря, то результаты исследований на выборке 8 лет были бы слишком оптимистичными. Таким образом, исключение из выборки данных лишь одного фрагмента, занимающего менее 5 % от общего объема, может, в худшем случае, существенно снизить оценку величины ошибки модели в худшем случае. Ошибка в 2015 году сильно превышает ошибку в любой другой год для всех моделей.

3. В условиях дефицита данных и отсутствия истории наблюдений за дополнительными факторами, такими как уровень снежного покрова в горах, число осадков, температура и т.д. нецелесообразно пытаться применять сложные модели машинного обучения. Хотя они могут извлекать скрытые зависимости даже из данных только самого временного ряда, но эти зависимости могут оказаться неустойчивыми во времени.

4. Регулярное и частое обучение модели на постоянно расширяющейся выборке данных может существенно повысить точность прогнозирования, поскольку изменения скрытых зависимостей целевой переменной от выходных признаков в этом случае максимально быстро попадают в обучающую выборку. Снижение ошибки для MLP составило в среднем 5.03 процентных пункта (16.5 %, 0.89 м³/с), для AdaBoost 2.35 п.п. (9.6 %, 0.43 м³/с), для Random Forest 2.66 п.п. (10.5 %, 0.48 м³/с). Для наиболее сложного 2015 года эффект еще сильнее, для MLP 11.83 п.п. (22.52 %, 2.13 м³/с), для AdaBoost 16.59 п.п. (36.05 %, 2.99 м³/с), для Random Forest 12.52 п.п. (26.23 %, 2.25 м³/с).

5. Ограничение размера модели может быть эффективным средством для борьбы с переобучением. Поскольку при фиксированном небольшом размере модель не имеет возможности выполнить слишком сильную подгонку под данные. В результате такие модели очень быстро обучаются на простейшем оборудовании и могут быть обучаемы в автономном режиме без контроля по стороны человека-эксперта. Кроме того, ансамблевые модели, строящиеся над деревьями решений, по своей природе более детерминированы, чем нейросетевые модели, поскольку работают на более простых жадных эвристиках, таких как ветвление для максимального снижения энтропии на каждом шаге формирования дерева.

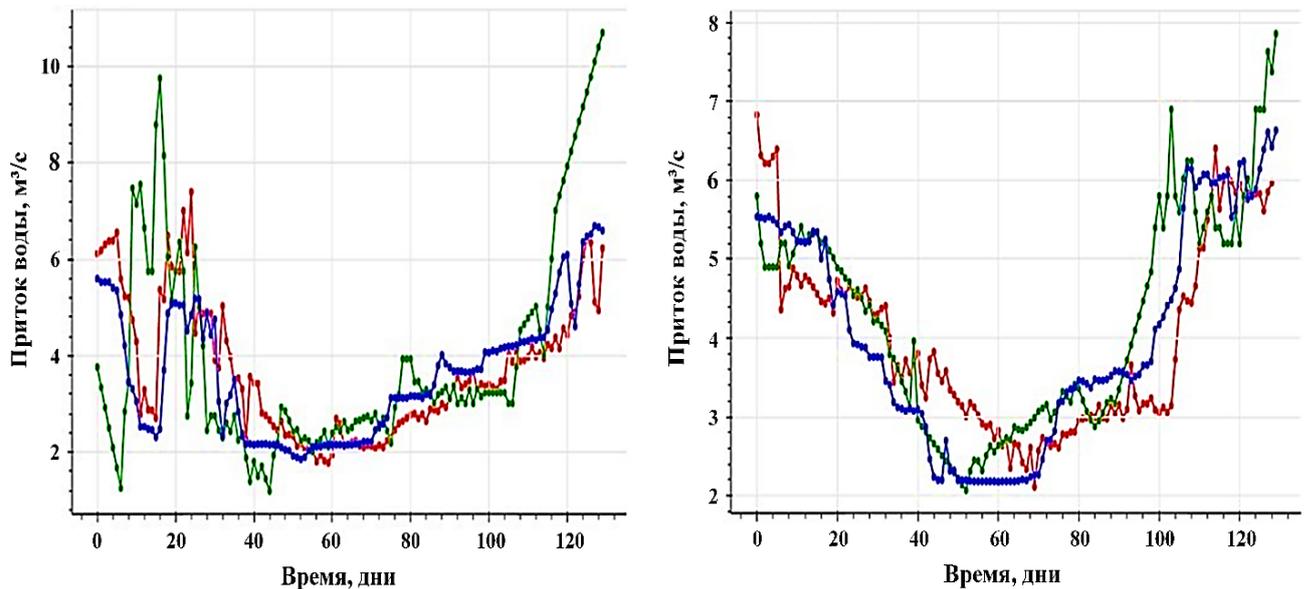


Рисунок 3.11 - Результаты для 2015 года (слева) 2019 (справа), зеленая линия – истинные значения, синяя – прогнозы АВ без самоадаптации (mRMSE 46.02 %, и 13.88 %), красная – с самоадаптацией (29.43 % и 12.62 %)

3.3 Среднесрочная модель прогнозирования генерации гидроэлектростанциями

Гидроэнергетика на сегодняшний день остается наиболее важной из возобновляемых источников энергии для производства электроэнергии во всем мире, обеспечивая 19% электроэнергии планеты [110,95]. Малая гидроэлектростанция (МГЭС) в большинстве случаев является “речной”, без

плотины или водохранилища и является одной из наиболее экономически эффективных и экологически безопасных энергетических технологий, которые следует рассматривать для электроснабжения сельских районов в менее экономически развитых странах [111,112,96,97].

Точное и надежное прогнозирование генерации ГЭС имеет важное значение для управления и планирования гидрогенерации. Ввиду бурного развития распределенной генерации недостаточно исторических данных о производстве электроэнергии, что затрудняет разработку прогнозных моделей. Задача прогнозирования генерации ГЭС считается одной из самых сложных задач в области интеллектуального анализа данных, требующая комплексного анализа больших объемов данных с учетом влияния множественных взаимосвязей и динамических процессов на электропотребление.

3.3.1 Постановка задачи и исходные данные

В данной работе рассмотрен подход по использованию компактных моделей на базе машинного обучения для среднесрочного прогнозирования генерации гидроэлектростанциями с учётом изменений температуры, отличием которого является полная автономность и автоматизация обучения и исключение риска переобучения.

Задача прогнозирования в данном разделе работы сформулирована следующим образом. Построение модели, выполняющей преобразование входных данных (признаков) X в прогноз вырабатываемой за сутки электроэнергии мощностей генерации на неделю (7 дней) вперед Y^* :

$$Y^* = f(X) \quad (3.8)$$

Показателем точности прогноза в данной работе выбран средний модуль процента ошибки (MAPE):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i^* - y_i}{y_i} \right| \quad (3.9)$$

где n – число дней в выборке; y_i^* – прогнозируемое значение генерируемой электроэнергии в i -е сутки; y_i – истинное значение в i -е сутки.

Используемый набор данных содержит суточные значения вырабатываемой электроэнергии (изолированная энергосистема ГБАО) с ноября по март включительно за 5 лет (2015–2019), и среднесуточные температуры, данные за 765 суток. Для обработки данных номера месяцев были преобразованы следующим образом: ноябрь – 1, декабрь – 2, ... март – 5. Отсчет лет начинается с 2015, так что номера годов преобразованы к диапазону 1–5. На Рисунке 1 приведены графики за один и тот же месяц (январь).

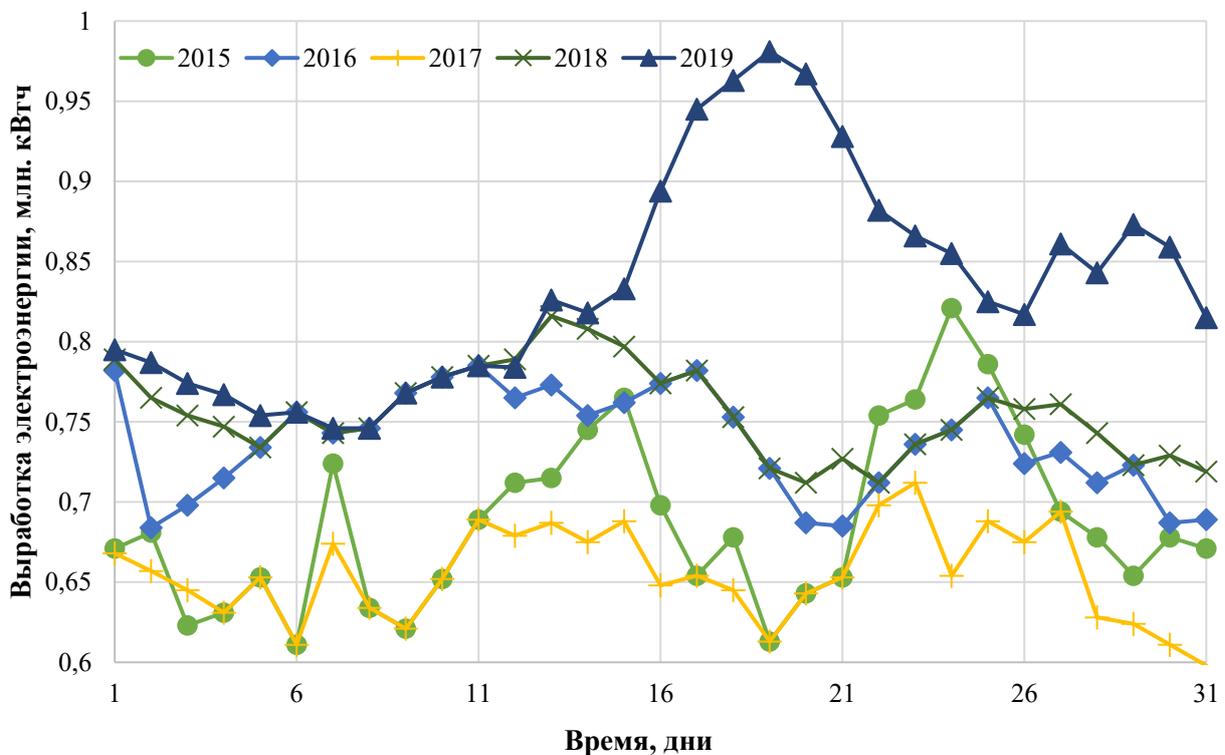


Рисунок 3.12 – Фрагмент выборки данных (январь 2015–2019 гг.)

3.3.2 Анализ данных и формирование выборки для применения моделей машинного обучения

Перед построением модели прогнозирования необходимо провести оценку и отбор наиболее значимых признаков, что может существенно повлиять на процесс обучения моделей на базе машинного обучения [113]. Помимо календарных

данных (год, месяц и число месяца), для прогноза необходимо использовать ретроспективные данные генерации. Также в работе проверяется гипотеза о влиянии среднесуточной температуры на точность прогноза.

На Рисунке 3.13 показаны матрица, состоящая из коэффициентов корреляции Пирсона для факторов, которые могут быть использованы при прогнозировании. Кроме них, используются ретроспективные данные генерации в предыдущие дни. Видно, что из метеорологических факторов существенное влияние оказывает температура (коэффициент 0,357), кроме того, явно наблюдается изменение год от года (коэффициент 0,375).

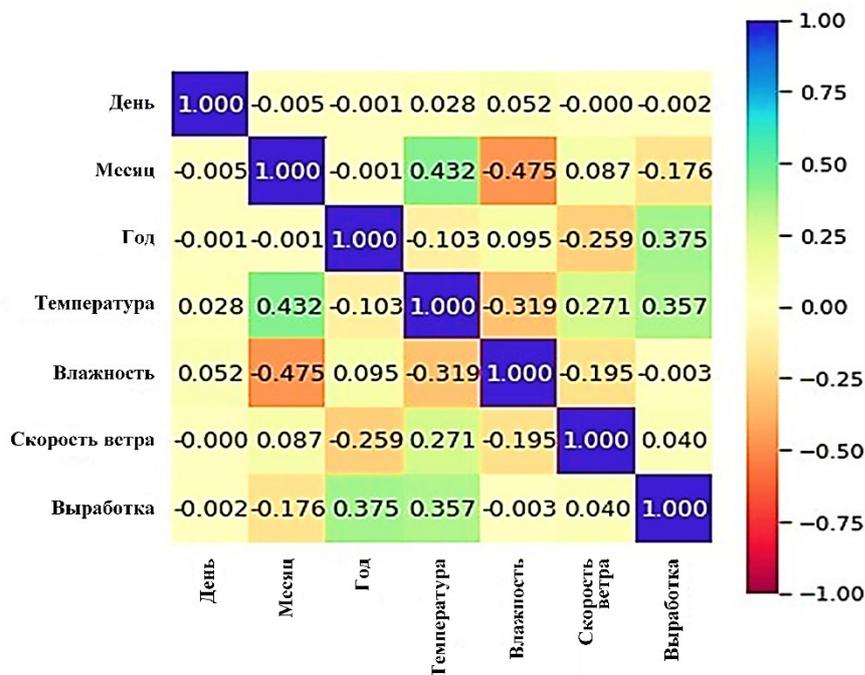


Рисунок 3.13 – Матрица коэффициентов корреляции

Сформированная в итоге структура выборки данных для построения регрессионных моделей показана в Таблице 1, где m – длина используемого временного ряда ретроспективных значений. Если используется неделя, от $m = 7$, две недели – 14 и т.д. Например, если $m = 7$ и дается прогноз на 26 января, то для формирования прогноза будут использованы ретроспективные данные с 19 ($\Gamma_7, 26 - 7 = 19$) по 13 ($\Gamma_{13}, 26 - [7 + m - 1] = 13$) января включительно.

Таблица 3.10 – Структура выборки для применения машинного обучения и пример заполнения

День	Месяц	Год	Среднесуточная температура, С ⁰	Г ₇ , млн. кВтч	Г ₈ , млн. кВтч	...	Г _{<i>m</i>+7-2} , млн. кВтч	Г _{<i>m</i>+7-1} , млн. кВтч	Г, млн. кВтч (прогнозируемое значение)
26	3	2	-3,2	0,687	0,721	...	0,684	0,782	0,724

Число записей в выборке зависит от m , так как для первых $m + 7$ дней начального месяца (ноября) нет нужного количества ретроспективных данных. Выборка делится на обучающе-валидационную и тестовую в пропорции 4 к 1 (в данном случае тестовую выборку попадает последний год всей выборки). Нужно отметить, что выборка не перемешивалась, так что тестирование было проведено в условиях, соответствующих реальным, когда модель обучается на ретроспективных данных, а затем используется в будущем на новых данных. В Таблице 3.11 приведены размеры выборок при разных значениях m .

Таблица 3.11 – Размеры выборок при разных интервалах используемых ретроспективных данных

m	Число столбцов	Число строк в обучающей выборке	Число строк в валидационной выборке	Число строк в тестовой выборке
7	11	576	110	137
14	18	547	104	130
21	25	518	98	123

3.3.3 Используемые модели машинного обучения и результаты

Модели, подбираемые гипер-параметры и диапазоны их значений приведены в Таблице 3.12 Для настройки гипер-параметров модели был применен случайный поиск (Random Search). Подбор гипер-параметров выполнялся при m равном 14. Программные реализации моделей, кроме XGBoost [114], взяты из открытой

библиотеки Scikit-Learn [108]. Для не приведенных в Таблице 3.12 гипер-параметров использованы значения по умолчанию из указанных источников.

Для всех моделей был проведен анализ влияния длины используемого временного ряда предыдущих значений генерации и влияние температуры на точность прогнозов.

Таблица 3.12 – Результаты подбора гипер-параметров примененных моделей машинного обучения

Модель	Гипер-параметр	Мин. значение	Макс. значение	Шаг	Подобранное значение
Линейная / полиномиальная регрессия с регуляризацией Тихонова [4*] (LR)	степень полинома	1	3	1	1
	к-ближайших соседей (kNN)	2	8	1	2
Адаптивный бустинг деревьев решений (ABDT)	число базовых моделей	10	50	1	14
	глубина деревьев	2	6	1	3
	минимальное число точек данных для ветвления	2	8	1	3
Адаптивный бустинг линейных моделей (ABLR)	число базовых моделей	2	10	1	2
Случайный лес (RF)	число базовых моделей	10	50	1	37
	глубина деревьев	2	10	1	5
	минимальное число точек данных для ветвления	2	8	1	2
Экстремальный градиентный бустинг (XGB)	число базовых моделей	10	50	1	29
	глубина деревьев	2	10	1	4
	скорость обучения	0.01	0.5	0.05	0.3
Многослойный перцептрон (MLP)	число нейронов первого скрытого слоя	10	200	10	90
	число нейронов первого скрытого слоя	10	100	10	80

Результаты моделей при различных вариантах используемых входных признаков на тестовой выборке приведены в Таблице 3.13 и на Рисунке 3.14.

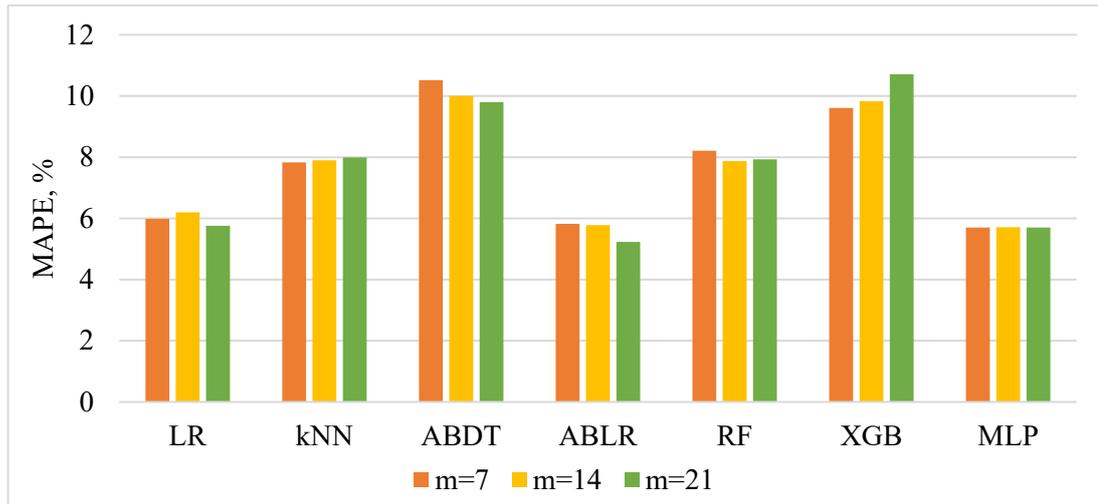


Рисунок 3.14 – Сравнение результатов моделей прогнозирования

Таблица 3.13 – Сравнение результатов моделей, приведены значения метрики ошибки MAPE, в каждой строке выделены наилучшие результаты, полученные для соответствующей модели

Модель	Без использования температуры			С использованием температуры			
	<i>m</i>	7	14	21	7	14	21
LR		5,99	6,20	5,76	6,53	7,15	7,19
kNN		7,83	7,90	7,99	13,41	14,12	13,35
ABDT		10,52	10,00	9,80	10,78	10,35	10,01
ABLR		5,82	5,78	5,23	6,25	7,30	6,98
RF		8,21	7,87	7,93	8,29	8,13	7,72
XGB		9,61	9,83	10,71	10,84	11,80	12,76
MLP		5,70	5,71	5,70	8,69	9,86	10,12

Из полученных результатов можно сделать следующие заключения:

1. Ансамбли деревьев решений, то есть моделей с дискретным (кусочно-постоянным) выходом уступает моделям с непрерывным выходом на 2-5 процентных пункта.

2. Для моделей с использованием линейной регрессии (LR и ABLR) увеличение интервала используемых ретроспективных данных несколько снижает ошибку, можно предположить, что за счет более точного определения тренда.

3. Наилучшая точность получена при использовании адаптивного бустинга с линейной регрессией в качестве базовой модели, что совпадает с результатами,

полученными ранее для прогнозирования электропотребления данной системы [115].

4. Существенное снижение точности моделей при использовании метеорологических данных является несколько неожиданным результатом, который требует дополнительного анализа. Вероятно, изменение климата год от года приводит и к изменению зависимостей между генерацией и температурой, так что модель, обученная на данных четырех первых лет, находит зависимости между температурой и генерацией, которые на следующий год оказываются измененными.

Это подтверждается визуализацией, показанной на Рисунках 3.15 - 3.17. Их анализ показывает, что зависимости генерации от температуры существенно отличаются из года в год. Даже коэффициенты корреляции существенно отличаются, в 2015 коэффициент корреляции Пирсона между генерацией и температурой составлял 0.38, в 2019 – 0.77, то есть в 2019 году температура оказывала существенно большее влияние на генерацию. Поэтому при обучении модели на ретроспективных данных она выделяет определенные зависимости, которые затем перестают работать из-за указанных выше причин.



Рисунок 3.15 – Распределение дней по генерации и температуре, 2015 год слева, 2016 справа

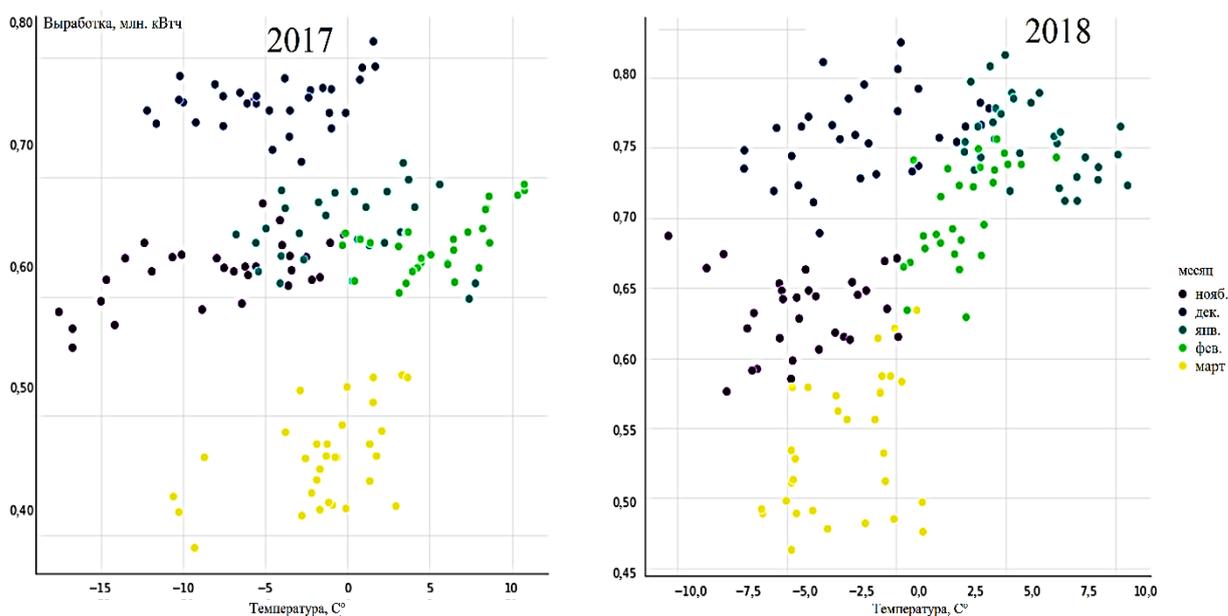


Рисунок 3.16 – Распределение дней по генерации и температуре, 2017 год слева, 2018 справа

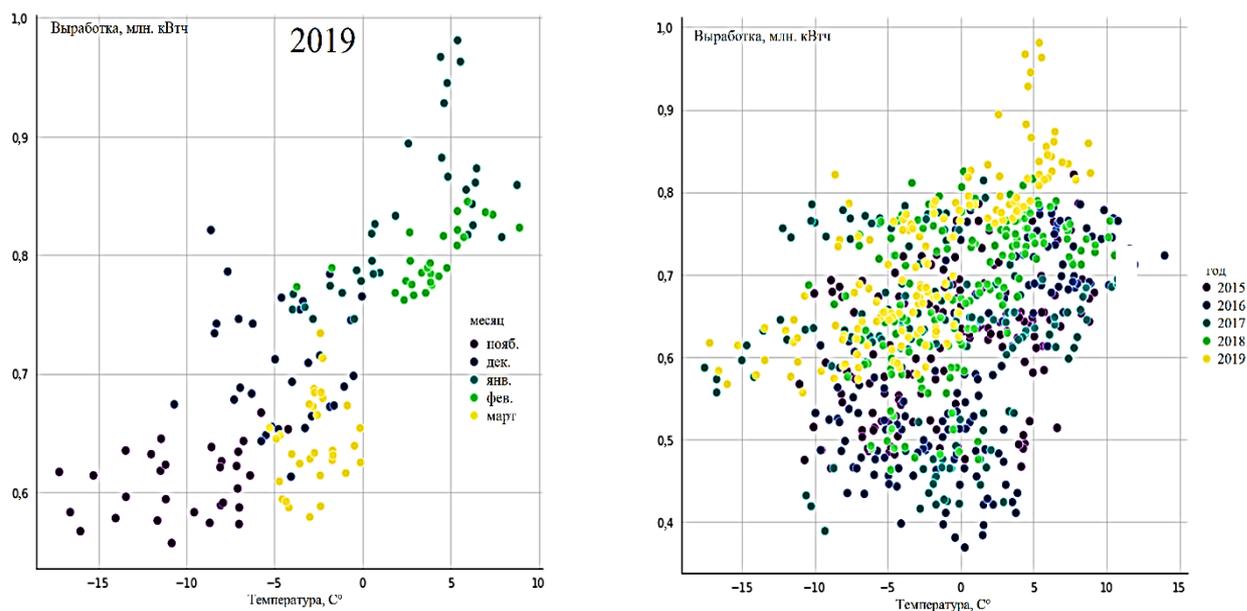


Рисунок 3.17 - Распределение дней по генерации и температуре, 2019 год слева, 2015-2019 гг. справа

Кроме того, из правой части Рисунка 3.17 видно, что 2019 год, который попал в тестовую выборку, в принципе характеризуется большим значением генерации, чем предыдущие 4 года, особенно при температурах выше $0\text{ }^{\circ}\text{C}$.

Сопоставление прогноза, полученного наилучшей для задачи моделью

(Адаптивный бустинг линейных регрессий с регуляризацией Тихонова) показано на Рисунке 3.18. Средняя ошибка на тестовой выборке составила 5,23 % или 44 МВтч при средней суточной генерации 718 МВтч.

Регуляризация Тихонова (L2-регуляризация) может быть записана следующим образом:

$$w^* = \arg \min_w \left(\sum_{i=1}^n (y_i - (wX_i - b))^2 + \lambda \|w\|_2^2 \right) \quad (3.10)$$

В выражении (3) w , b – коэффициенты регрессии, $\|w\|_2^2$ – квадрат Евклидова расстояния или сумма квадратов всех элементов вектора w .

Поскольку модель адаптивного бустинга представляет собой линейную взвешенную сумму базовых моделей, то итоговая ансамблевая модель, объединяющая 4 линейные регрессии, может быть после завершения обучения свернута с линейную форму

$$Y^* = f(X) = AX + b$$

$$y^* = f(x_i) = a_1 year_i + a_2 month_i + a_3 day_i + a_4 E_{7,i} + a_5 E_{8,i} + \dots + a_{25} E_{28,i} + b$$

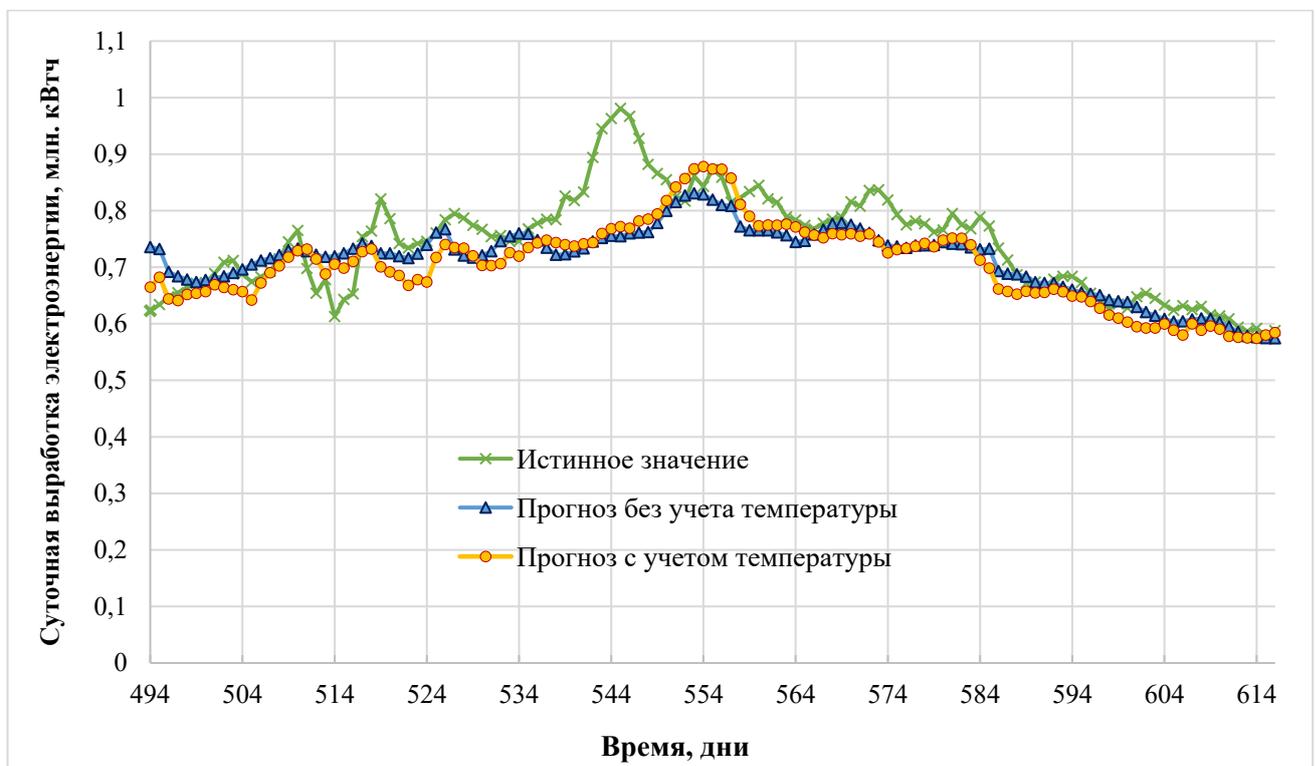


Рисунок 3.18 - Сопоставление прогнозов и истинных значений мощности, 2019 г. (тестовая часть выборки)

Выводы по главе 3

В рамках третьей главы диссертационного исследования основное внимание было уделено проблеме среднесрочного прогнозирования (на неделю вперед) электропотребления, притока воды и генерации ИЭС ГБАО в холодное время года (ноябрь-март), когда наблюдается дефицит электроэнергии, вырабатываемой ГЭС, на основе компактных моделей машинного обучения.

Чтобы повысить точность прогнозирования электропотребления в работе использованы метеорологические факторы и кластеризация условий. Используются только ретроспективные данные, чтобы исключить зависимость от метеорологических прогнозов. На этапе предобработки данных применен метод разложения пространства признаков на главные компоненты, который, во-первых, позволяет проверить корректность кластеризации, во-вторых, агрегировать метеорологические признаки. В результате средняя ошибка прогноза на тестовой части выборки данных составила 0.59 МВт, среднеквадратическая ошибка 0.77 МВт, средняя ошибка в процентах 4.5 %. По сравнению с авто регрессионной моделью без кластеризации и учета метеорологических факторов (MAE 0.80 МВт, RMSE 0.96 МВт, MAPE 6.11 %).

Для решения задачи среднесрочного прогнозирования притока воды в условиях дефицита достоверной информации в работе предложено и обосновано использование компактных моделей на основе ансамблей неглубоких деревьев решений. Ошибка (RMSE) ансамблевых моделей за счет самоадаптации сократилась с 4.4–4.5 м³/с до 4.0 м³/с и оказалась ниже ошибки более сложной многослойной рекуррентной модели (4.9 м³/с). В самый трудный для прогнозирования год за счет непрерывной самоадаптации ошибка ансамблевых моделей сократилась с 8.3–8.6 м³/с до 5.3–6.3 м³/с.

Выполнен анализ семи различных моделей для прогнозирования генерации гидроэлектростанциями на основе ретроспективных данных о генерации и метеорологических параметров. Наилучший результат был получен при использовании адаптивного бустинга с линейной регрессией в качестве базовой модели. Средняя ошибка на тестовой выборке составила 5,23 % или 44 МВтч при средней суточной генерации 718 МВтч.

ГЛАВА 4 КОМПЛЕКСНОЕ ПЛАНИРОВАНИЕ ГРАФИКА ГЕНЕРАЦИИ НА СРЕДНЕСРОЧНОЮ ПЕРСПЕКТИВУ

4.1 Планирование генерации

На протяжении многих лет население мира быстро росло, что привело к росту индустриализации, что в конечном итоге потребовало увеличения генерации. Существует два вида энергетических ресурсов: традиционные и нетрадиционные. Традиционные энергетические ресурсы ограничены и вызывают загрязнение окружающей среды, в то время как нетрадиционные возобновляемые источники энергии являются экологически более чистыми. Среди всех ресурсов вода является эффективной, чистой и используется гидроэлектростанциями для генерации. Ежегодная выработка электроэнергии ГЭС зависит от управления годовым сбросом при планировании работы ГЭС и ее водохранилища. Планирование гидроэнергетических систем является сложным, поскольку оно требует обеспечения покрытия пикового спроса на электроэнергию в течение всего рабочего процесса системы. При правильном планировании при том же количестве воды, поступающей в водохранилище, вырабатывается больше энергии, что обеспечивает значительную экономию для производителя даже при незначительном улучшении вычислений [116,117].

Планирование электроэнергетической системы разделено в соответствии с временным горизонтом, т.е. на день, неделю или год. Еженедельное планирование известно как среднесрочное планирование. Годовое планирование считается долгосрочным планированием. Подход к долгосрочной оптимизации обычно следует за горизонтом планирования в несколько лет [118,119]. Подход к краткосрочной оптимизации основан на временном горизонте от нескольких часов до суток.

Планирование генерации является одним из основных аспектов производства электроэнергии за счет снижения общей стоимости вырабатываемой электроэнергии. Для решения задачи планирования генерации используются

различные подходы и методы; смешанное целочисленное программирование [120,121], динамическое программирование [122,125], иерархическая оптимизация [126], метод релаксации Лагранжа [127,128], поиск с запретами [129] и методы нелинейного программирования [130].

В настоящей работе задача планирования графика генерации решается на основе прогнозирования электропотребления и генерация с учетом метеорологических факторов, для того чтобы уменьшить использование ДЭС за счёт повышения точности прогнозирования и оптимизации регулирования выработки ГЭС.

4.2 Анализ и предварительная обработка данных

Использован набор данных суточных значений электропотребления ГБАО и выработки ГЭС всей области за период ноябрь-март 2015-2020 гг., суммарно 30 месяцев или 908 суток. Рассмотрены только эти месяцы, так как именно в этом сезоне наблюдается дефицит электроэнергии, вырабатываемой ГЭС, который компенсируется за счет ДЭС.

На Рисунках 4.1- 4.5 показаны графики электропотребления и выработки ГЭС и ДЭС. Видно, что график выработки ГЭС намного менее предсказуем, чем график электропотребления. Но выработкой ГЭС можно управлять с помощью регулирования сбросов воды из водохранилищ.

Для прогнозирования потребления, выработки и их разности был сформирован набор данных, структура которого показана в Таблице 4.1. Для среднесрочного планирования использования ДЭС нужны среднесрочные прогнозы потребления и выработки ГЭС, разность которых и дает необходимую выработку ДЭС. Но также можно строить сразу прогноз разности потребления и выработки ГЭС, поэтому в набор данных добавлен столбец *delta* (Δ).

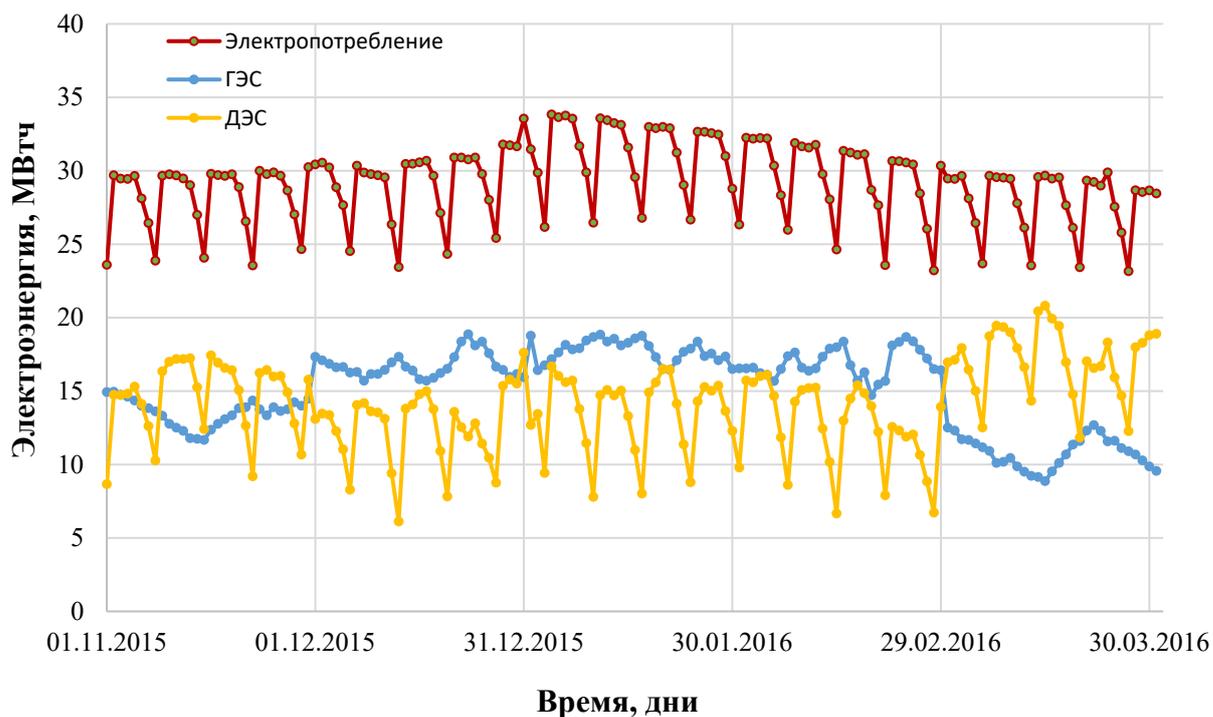


Рисунок 4.1 – Графики электропотребления и генерации ноябрь 2015 – март 2016

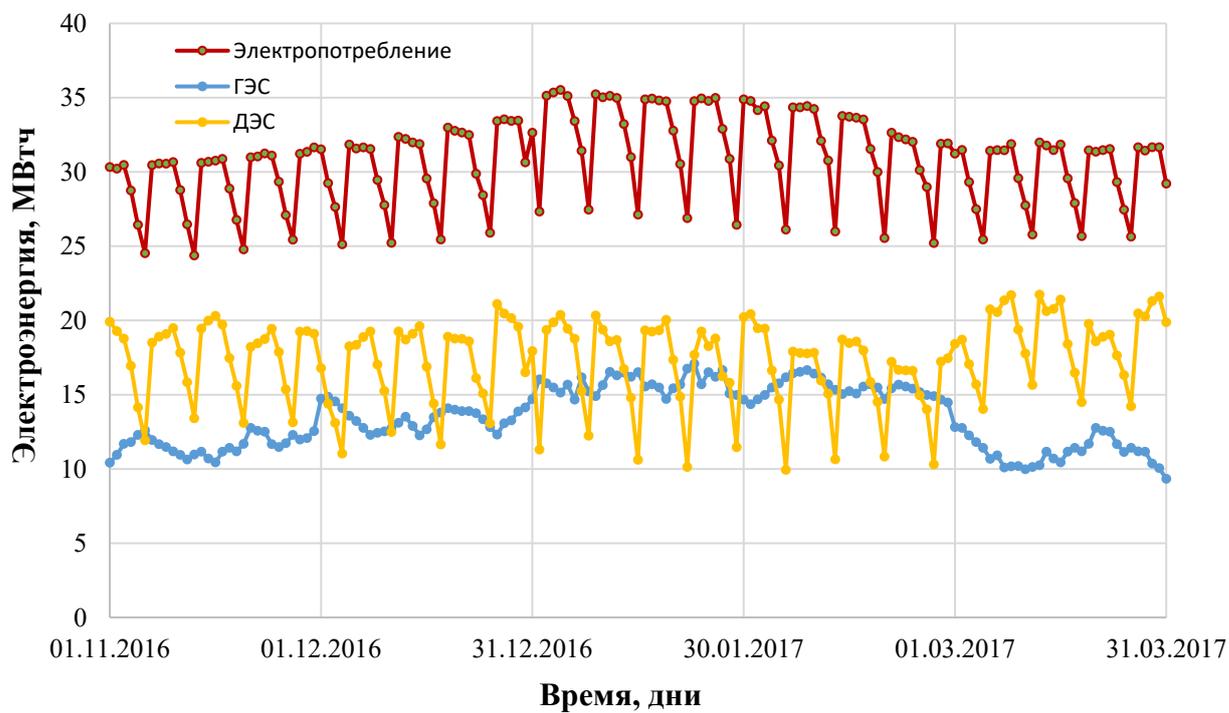


Рисунок 4.2 – Графики электропотребления и генерации ноябрь 2016 – март 2017

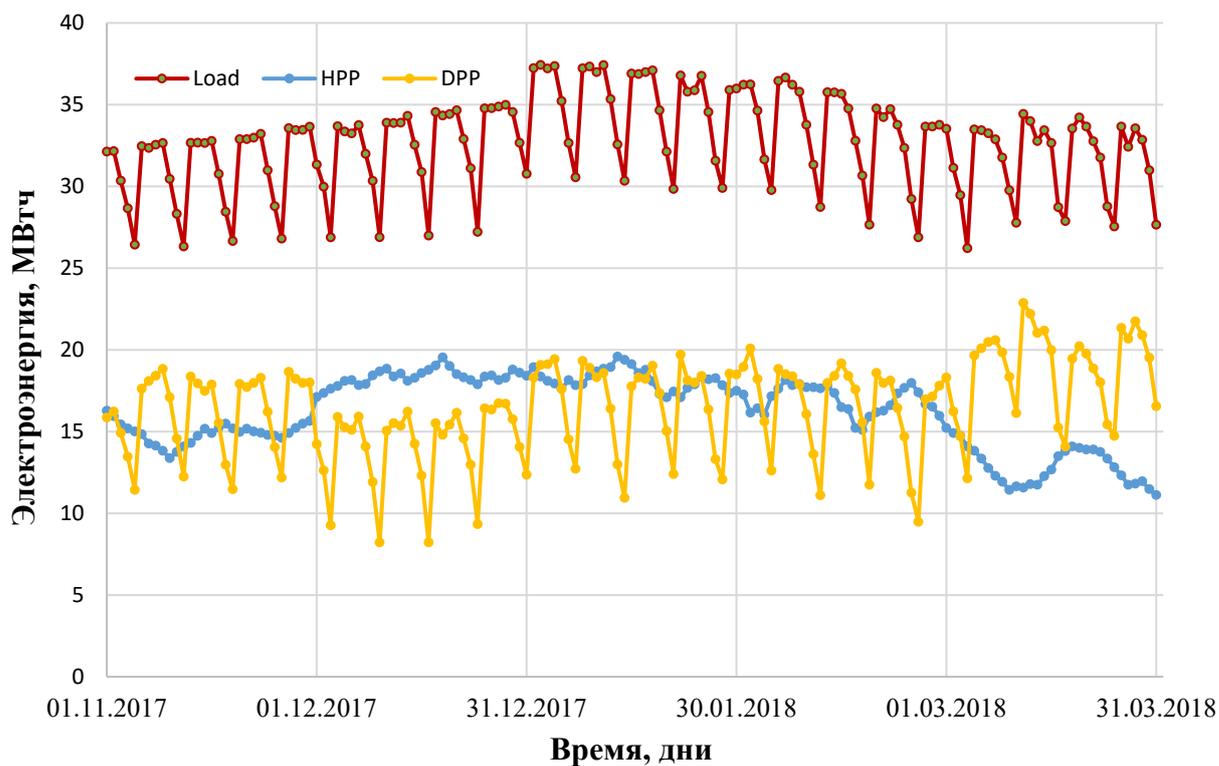


Рисунок 4.3 – Графики электропотребления и генерации ноябрь 2017 – март 2018

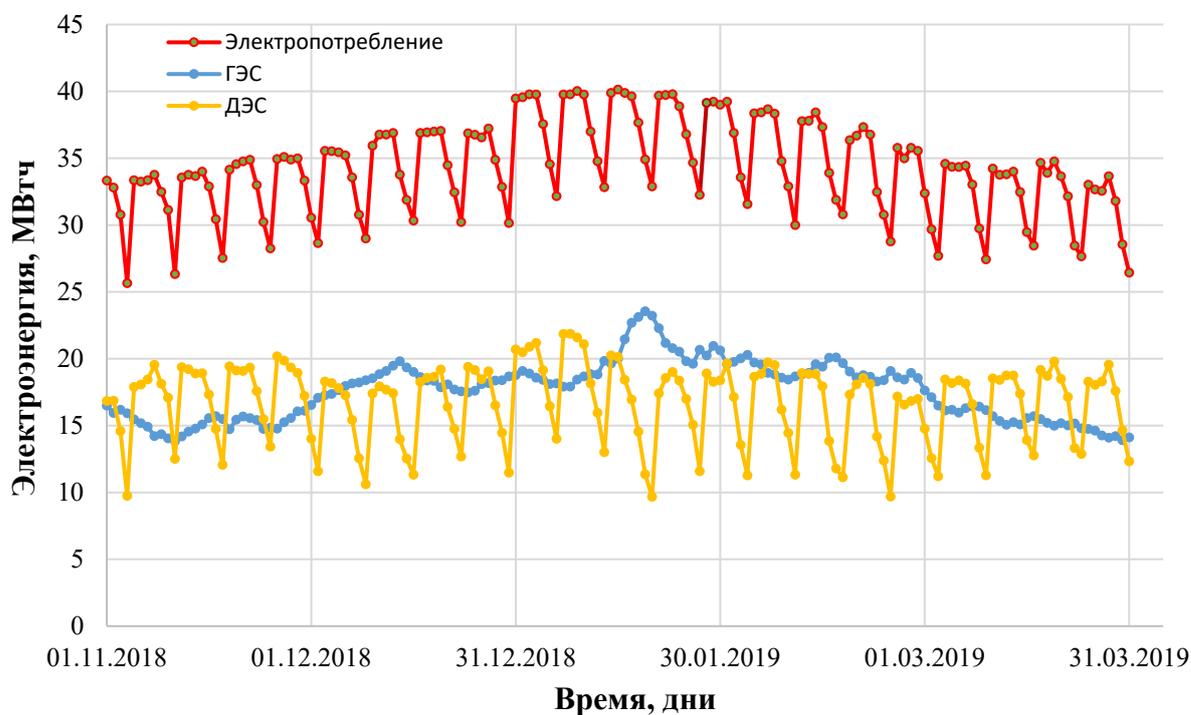


Рисунок 4.4 – Графики электропотребления и генерации ноябрь 2018 – март 2019

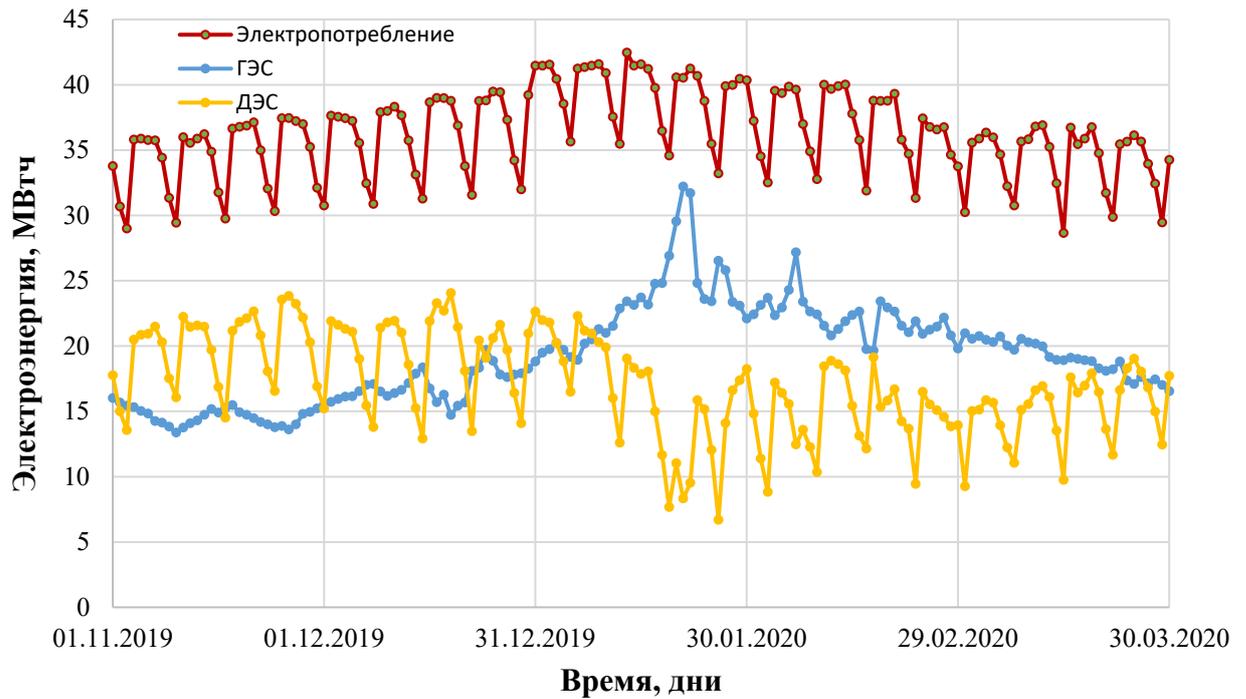


Рисунок 4.5 – Графики электропотребления и генерации ноябрь 2019 – март 2020

Таблица 4.1 – Структура датасета, отмеченные * признаки исключены после анализа

Имя столбца	Тип столбца	Пояснение	Единица измерения
год	признак	год от 2015	-
месяц	признак	номер месяца в году	-
День	признак*	число месяца	-
температура	признак	прогноз температуры воздуха	С ⁰
влажность	признак*	прогноз влажности воздуха	%
скорость ветра	признак	прогноз скорости ветра	м/с
облачность	признак*	прогноз уровня облачности	%
день недели	признак	рабочий или выходной день	-
электропотребления	цель	электропотребление за сутки	МВтч
выработка ГЭС	цель	выработка ГЭС за сутки	МВтч
delta (Δ)	цель	электропотребления – выработка ГЭС	МВтч

На Рисунке 4.6 показана матрица коэффициентов корреляции Спирмена. Из матрицы видно, что на потребление наибольшее влияние оказывают тип дня (рабочий или выходной), скорость ветра, температура и год, поскольку наблюдается тренд повышения электропотребления. То же самое справедливо для выработки ГЭС, кроме корреляции с типом дня. Номер дня в месяце, влажность и облачность влияния на точность прогноза не оказывают, что подтвердилось в ходе предварительных экспериментов, поэтому они были исключены из рассмотрения.

Год	1,00	0,00	0,00	-0,09	0,04	-0,35	-0,06	0,00	0,70	0,60	0,27
Месяц	0,00	1,00	0,01	0,39	-0,41	0,08	-0,06	0,00	-0,09	-0,20	0,10
День	0,00	0,01	1,00	0,00	0,04	0,02	0,05	-0,01	0,01	-0,01	0,01
Температура	-0,09	0,39	0,00	1,00	-0,38	0,27	0,16	-0,01	-0,33	-0,52	0,08
Влажность	0,04	-0,41	0,04	-0,38	1,00	-0,16	0,55	-0,04	0,20	0,19	0,08
Скорость ветра	-0,35	0,08	0,02	0,27	-0,16	1,00	0,15	0,05	-0,37	-0,42	-0,03
Облачность	-0,06	-0,06	0,05	0,16	0,55	0,15	1,00	-0,04	0,01	-0,07	0,10
День недели	0,00	0,00	-0,01	-0,01	-0,04	0,05	-0,04	1,00	-0,51	0,00	-0,63
Нагрузка	0,70	-0,09	0,01	-0,33	0,20	-0,37	0,01	-0,51	1,00	0,62	0,61
Выработка ГЭС	0,60	-0,20	-0,01	-0,52	0,19	-0,42	-0,07	0,00	0,62	1,00	-0,18
Δ	0,27	0,10	0,01	0,08	0,08	-0,03	0,01	-0,63	0,61	-0,18	1,00
	Год	Месяц	День	Температура	Влажность	Скорость ветра	Облачность	День недели	Нагрузка	Выработка ГЭС	Δ

Рисунок 4.6 – Коэффициенты корреляции Спирмена

Исходный набор данных может быть записан как

$$\{(T_i, W_i, E_i, H_i, \Delta_i,)\}, i = 1, \dots, n. \quad (4.1)$$

где i – номер суток в выборке; T – вектор календарных признаков (*месяц, год, неделя_день*); W – вектор метеорологических признаков; E – электропотребление; H – выработка ГЭС; Δ – разность E и H :

$$\Delta = E - H \quad (4.2)$$

На этапе предобработки набор данных преобразуется в три набора данных, соответствующих трем задачам прогнозирования (потребления, выработки ГЭС, разности потребления и выработки ГЭС):

$$\{(T_i, W_i, E_{i-w-h}, E_{i-w+1-h}, \dots, E_{i-1-h}, E_{i-h}, E_i,)\}, i = 1, \dots, n. \quad (4.3)$$

$$\{(T_i, W_i, H_{i-w-h}, H_{i-w+1-h}, \dots, H_{i-1-h}, H_{i-h}, H_i,)\}, i = 1, \dots, n. \quad (4.4)$$

$$\{(T_i, W_i, \Delta_{i-w-h}, \Delta_{i-w+1-h}, \dots, \Delta_{i-1-h}, \Delta_{i-h}, \Delta_i,)\}, i = 1, \dots, n. \quad (4.5)$$

где w определяет количество предыдущих суток, которые будут использоваться для построения прогноза, h – горизонт прогнозирования (1 неделя, так что $h = 7$). Количество предыдущих суток w в ходе предварительных экспериментов также выбрано равным 7, что отражает недельный цикл процессов в энергосистеме [131].

Кроме того, на этапе предобработки данных выполняется Min-Max нормализация, чтобы все значения признаков оказались в диапазоне от 0 до 1.

Датасет разделен на обучающую часть (2015-2019 гг.) и тестовую (2020 г.), то есть в пропорции 5 к 1.

4.3 Прогнозирование и критерии оптимизации

В соответствии с выражениями (4.3) - (4.5) возникает три однотипные формулировки задачи прогнозирования:

$$E_i^* = f_E(T_i, W_i, E_{i-w-1-h}, \dots, E_{i-h}) \quad (4.6)$$

$$\sum_i |E_i^* - E_i| \rightarrow \min \quad (4.7)$$

$$H_i^* = f_H(T_i, W_i, H_{i-w-1-h}, \dots, H_{i-h}) \quad (4.8)$$

$$\sum_i |H_i^* - H_i| \rightarrow \min \quad (4.9)$$

$$\Delta_i^* = f_D(T_i, W_i, \Delta_{i-w-1-h}, \dots, \Delta_{i-h}) \quad (4.10)$$

$$\sum_i |\Delta_i^* - \Delta_i| \rightarrow \min \quad (4.11)$$

Так как итоговая цель – планирование генерации ДЭС, то фактически первые две задачи нужны для следующего выражения:

$$\Delta_i^{**} = E_i^* - H_i^* \quad (4.12)$$

$$\sum_i |\Delta_i^{**} - \Delta_i| \rightarrow \min \quad (4.13)$$

При этом, как указано в предыдущем разделе работы, в описании задачи сокращения ДЭС за счет повышения точности планирования, конечной целью является не точный прогноз, а сокращение избыточного использования ДЭС на его основе. Уравнение суточного баланса генерации и потребления можно записать как (здесь и далее рассматриваются суточные значения, индекс суток i опущен для упрощения восприятия):

$$H + D - E = 0 \quad (4.14)$$

где H – выработка ГЭС, D – выработка ДЭС, E – потребление.

Потребление E рассматривается как нерегулируемая величина, а выработка ГЭС может корректироваться для обеспечения баланса [132]. Выработка ДЭС соответствует запланированной:

$$D = \Delta^* \quad (4.15)$$

или, в зависимости от метода прогнозирования:

$$D = \Delta^{**} \quad (4.16)$$

Из-за погрешности прогноза эта величина не будет равна разности H и E , которые представлены в датасете. Поэтому уравнение баланса принимает вид:

$$H - H_\delta + D - E = 0 \quad (4.17)$$

где H_δ – необходимое измерение выработки ГЭС по сравнению с запланированным.

Если $H_\delta > 0$, это означает, что от ГЭС в рассматриваемые сутки можно было взять больше электроэнергии, а от ДЭС меньше, следовательно, ошибка планирования привела к лишнему расходу ДЭС.

Если $H_\delta < 0$, это означает, что от ГЭС в данные сутки необходимо взять больше, чем запланировано. Но возможность увеличения выработки ГЭС ограничено как в отдельные сутки, так и суммарно длительном интервале времени.

Из выражений (4.2) и (4.17) получаем:

$$-\Delta - H_{\delta} + D = 0 \quad (4.18)$$

$$H_{\delta} = D - \Delta \quad (4.19)$$

это означает, что ошибка планирования ДЭС равна регулированию выработки ГЭС, как и заложено в начальной модели энергосистемы.

H_{δ} можно представить как

$$H_{\delta} = H_{\delta+} + H_{\delta-}, H_{\delta+} \geq 0, H_{\delta-} \leq 0 \quad (4.20)$$

Задача оптимизации сводится к минимизации излишнего потребления ДЭС $H_{\delta+} \rightarrow \min$, при этом есть ограничения на максимальное по модулю значение $H_{\delta-}$, а разность суммарных значения $H_{\delta+}$ и суммы модулей $H_{\delta-}$ в течение недели не может быть выше нуля, иначе будет нарушен водно-энергетический режим работы ГЭС, обусловленный ограниченной емкостью водохранилища.

В настоящей разделе работы оптимизация рассматривается в упрощенном виде, только на основе прогнозирования выработки ГЭС и потребления или прогнозирования их разности напрямую. Эксперименты показали, что при получаемой точности прогноза, которая показана в следующем разделе, указанные ограничения по регулированию ГЭС выполняются в большинстве случаев автоматически.

В Таблице 4.2 приведены результаты прогнозирования, полученные на тестовой выборке различными моделями:

- LR – линейная регрессия;
- ANN – многослойная нейронная сеть (многослойный перцептрон, архитектура 16-16-8-1) [133];
- RF – Random Forest;
- AB – AdaBoost (на основе деревьев решений);
- SVR – Support Vector Regression.

В Таблице 4.2 для каждой из моделей приведены результаты с использованием и без использования метеорологических прогнозов. Прогнозы на неделю вперед имеют не очень высокую точность для рассматриваемой энергосистемы и их использование может оказывать даже негативное влияние на

точность прогноза, что связано еще и с изменением климата, так как модели, построенные на ретроспективных данных, могут выделить зависимости, которые перестанут быть актуальными на следующий год [134].

Таблица 4.2 – Результаты прогнозирования

Модель	Использование метеофакторов	E^* , MAE, МВтч	E^* , nMAE, %	H^* , MAE, МВтч	H^* , nMAE, %	Δ^* , MAE, МВтч	Δ^* , nMAE, %	Δ^{**} , MAE, МВтч	Δ^{**} , nMAE, %
LR	да	0.643	1.729	1.528	7.413	1.705	10.296	1.640	9.906
LR	нет	0.692	1.862	1.491	7.230	1.583	9.558	1.614	9.750
ANN	да	0.696	1.872	1.367	6.632	1.502	9.072	1.475	8.913
ANN	нет	0.786	2.114	1.208	5.860	1.567	9.465	1.687	10.190
RF	да	0.952	2.562	2.674	12.969	1.658	10.015	2.674	12.969
RF	нет	0.942	2.533	2.182	10.586	1.603	9.678	1.913	11.556
AB	да	1.236	3.324	2.332	11.315	1.505	9.091	2.101	12.691
AB	нет	1.305	3.509	2.597	12.599	1.542	9.311	2.338	14.121
SVR	да	0.987	2.654	1.158	5.616	1.741	10.517	1.535	9.270
SVR	нет	1.056	2.841	1.418	6.8781	1.627	9.825	1.664	10.047

Ошибки для Δ^* и Δ^{**} – это ошибки прогнозов одной и той же величины Δ , полученных разными способами, как показано в выражениях (4.10), (4.12). Следует отметить, что из полученных результатов нельзя дать единой рекомендации о том, строить ли по-отдельности прогноз потребления и прогноз генерации или же непосредственно прогнозировать разность потребления и генерации ГЭС. Для ансамблевых моделей лучшей стратегией оказался первый вариант, для остальных нет выраженного преимущества той или иной стратегии.

Для каждого из прогнозируемых параметров в Таблице 4.2 зеленым цветом выделена строка, соответствующая модели с наименьшей ошибкой прогноза (при близких результатах выделено по две строки). Видно, что для большинства случаев нейронная сеть показала наилучший результат. Но в данной работе важен прогноз нехватки выработки ГЭС, так как на его основе планируется использование ДЭС. Для наглядности точности этого прогноза приведены в Таблице 4.3, при этом для каждой модели выбран наилучший сценарий между Δ^* и Δ^{**} .

Таблица 4.3 – Наилучшие результаты прогнозирования разности E и H

Модель	Ошибка прогноза разности E и H	
	MAE, МВтч	nMAE, %
LR	1.583	9.558
ANN	1.475	8.913
RF	1.603	9.678
AB	1.505	9.091
SVR	1.535	9.270

4.4 Оценка влияния прогноза на планирование генерации ДЭС

После получения прогнозов можно определить величины $H_{\delta+}$ и $H_{\delta-}$ по выражениям (4.19), (4.20). Результаты приведены в Таблице 4.4. В ней рассмотрены суммарные величины $H_{\delta+}$ и $H_{\delta-}$ за тестовый период (ноябрь 2019 – март 2020):

$$S_{\delta+} = \sum_{i=1}^m H_{\delta+} \quad (4.21)$$

$$S_{\delta-} = \sum_{i=1}^m H_{\delta-} \quad (4.22)$$

$$S_{\delta+} = -S_{\delta-} - S_{\delta+} \quad (4.23)$$

В Таблице 4.5 для наглядности отдельно указано избыточное использование ДЭС для наилучшего сценария использования прогнозов каждой модели.

Таблица 4.4 – Влияние ошибок прогноза на ДЭС

Модель	Использование метеофакторов	Δ^*				Δ^{**}			
		MAE, МВтч	$S_{\delta+}$, МВтч	$S_{\delta-}$, МВтч	ΔS , МВт	MAE, МВтч	$S_{\delta+}$, МВтч	$S_{\delta-}$, МВтч	ΔS , МВт
LR	да	1.705	109	-125	16	1.640	219	-48	-171
LR	нет	1.583	183	-58	-125	1.614	152	-93	-59
ANN	да	1.502	170	-59	-111	1.475	172	-53	-119
ANN	нет	1.567	178	-60	-118	1.687	208	-48	-160
RF	да	1.658	119	-133	14	2.674	341	-30	-311
RF	нет	1.603	120	-123	3	1.913	258	-33	-225
AB	да	1.505	116	-113	-3	2.101	278	-41	-237
AB	нет	1.542	115	-116	1	2.338	281	-47	-234
SVR	да	1.741	216	-49	-167	1.535	126	-107	-19
SVR	нет	1.627	188	-60	-128	1.664	50	-204	154

Таблица 4.5 – Наилучшие и наихудшие результаты каждой из моделей

Модель	Использование метеофакторов	лучшее $S\delta+$, МВтч	худшее. $S\delta+$, МВтч	отклонение лучшего от лучшего среди всех моделей	отклонение худшего от лучшего среди всех моделей
LR	да	109	219	0	110
LR	нет	152	183	43	74
ANN	да	170	172	61	63
ANN	нет	178	208	69	99
RF	да	119	341	10	232
RF	нет	120	258	11	149
AB	да	116	278	7	169
AB	нет	115	281	6	172
SVR	да	126	216	17	107

Наименьшее избыточное использование дизельного топлива оказалось у ансамблевой модели AdaBoost. Ансамблевая модель показала наиболее сбалансированный результат с точки зрения распределения видов ошибок, поэтому планирование на ее основе привело бы к наименьшему излишнему использованию ДЭС. Рисунок 4.11 показывает средние ошибки прогнозов и избыточное использование ДЭС. Результаты ANN с точки зрения использования ДЭС оказались хуже, что связано с перекосом ошибок в сторону занижения разности потребления и выработки ГЭС. Хотя на обучающей части датасета все модели обучались при равных значимостях обоих видов ошибок.

В то же время задать во время обучения большую значимость ошибок, которые ведут к снижению разности $E - H$, можно, но в этом случае возникает два риска. Во-первых, что в какие сутки прогноз окажется настолько занижен, что увеличение выработки ГЭС на нужную величину окажется невозможным. Во-вторых, что суммарная выработка ГЭС на каком-либо интервале времени приведет к недопустимому снижению уровня воды в водохранилище. В обоих случаях итогом будет невозможность обеспечить электроэнергией потребителей в полной мере. Например, планирование на основе SVR без использования метеорологических факторов привело бы к такому результату, поскольку ΔS составляет 154 МВтч за холодный сезон.

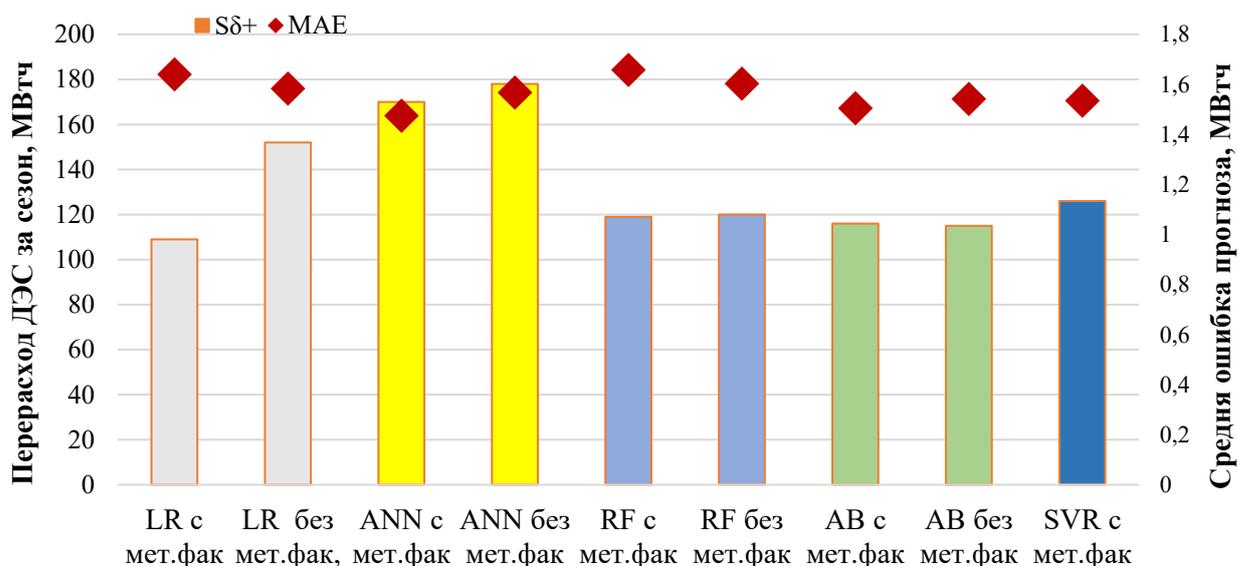


Рисунок 4.7 – Точность прогноза разности потребления и выработки ГЭС и избыточное потребление ДЭС в тестовый период (ноябрь 2019 – март 2020)

На Рисунках 4.8 – 4.12 показаны результаты прогнозов двух моделей: AbaBoost и ANN.

Видно, что ансамблевая модель имеет невысокую точность прогнозирования как потребления, так и выработки ГЭС. Но прогноз именно разности этих величин уже намного точнее и близок по точности к прогнозу нейросетевой модели (Рисунок 4.8), а главное – они более сбалансированы.

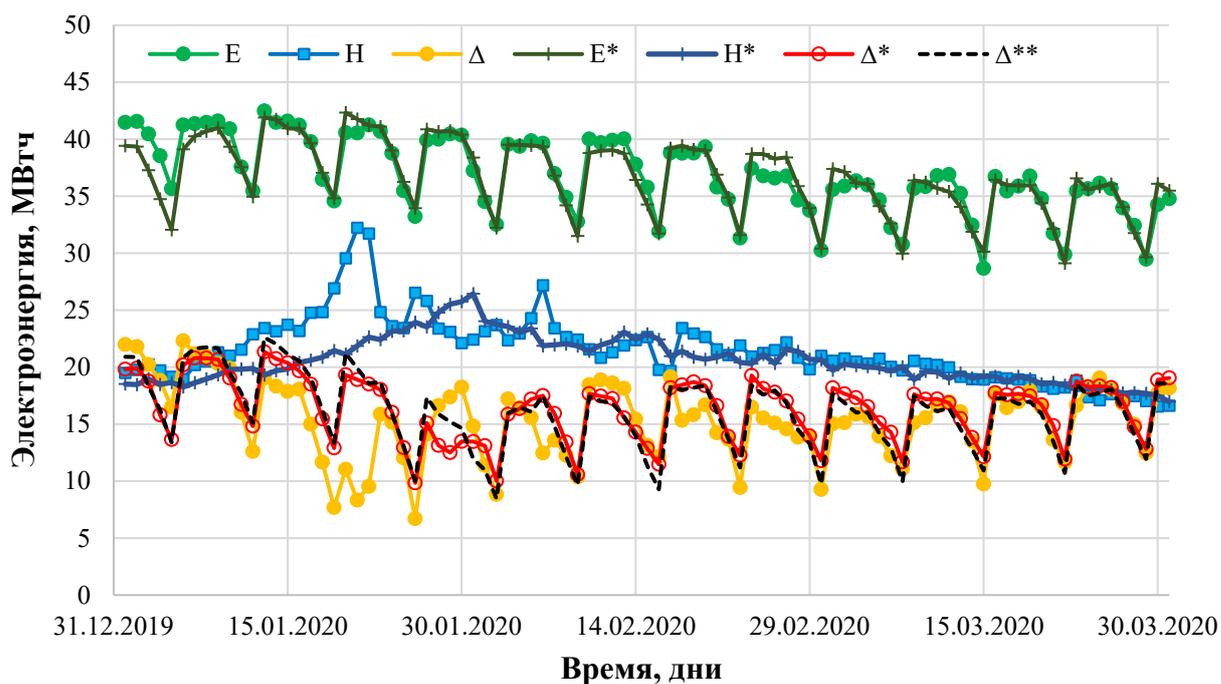


Рисунок 4.8 – Результаты прогноза ANN, январь – март 2020 г.

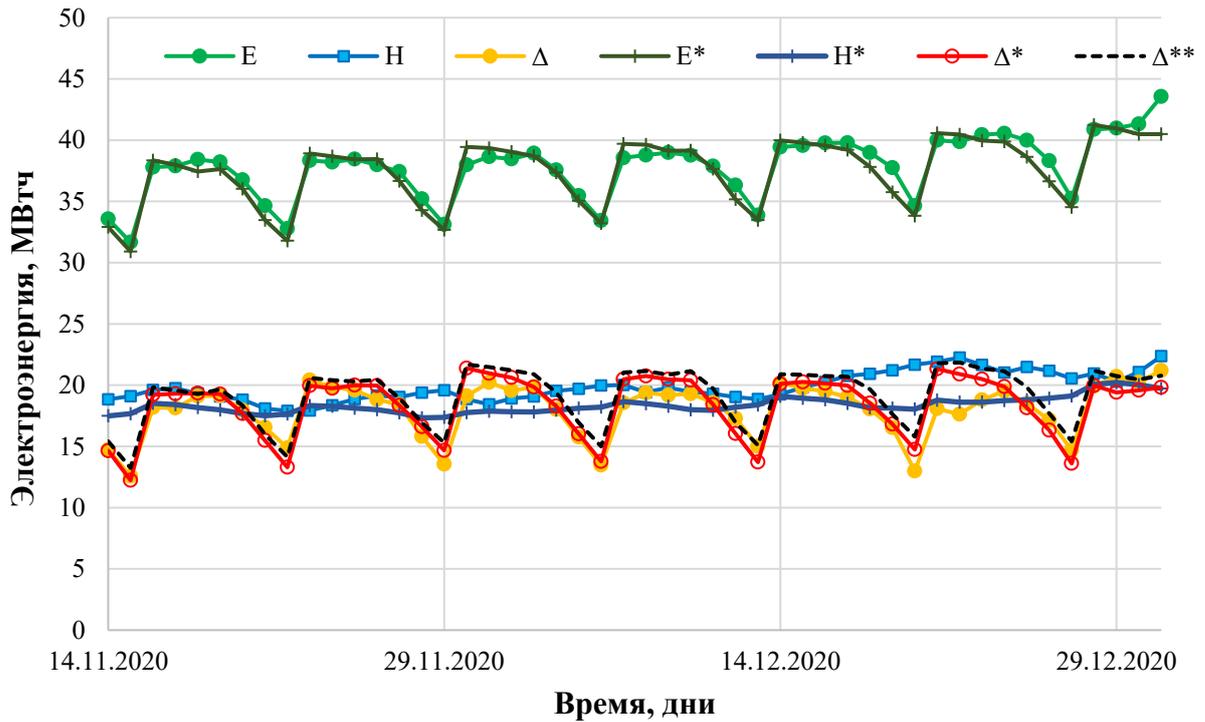


Рисунок 4.9 – Результаты прогноза ANN, ноябрь – декабрь 2020 г.

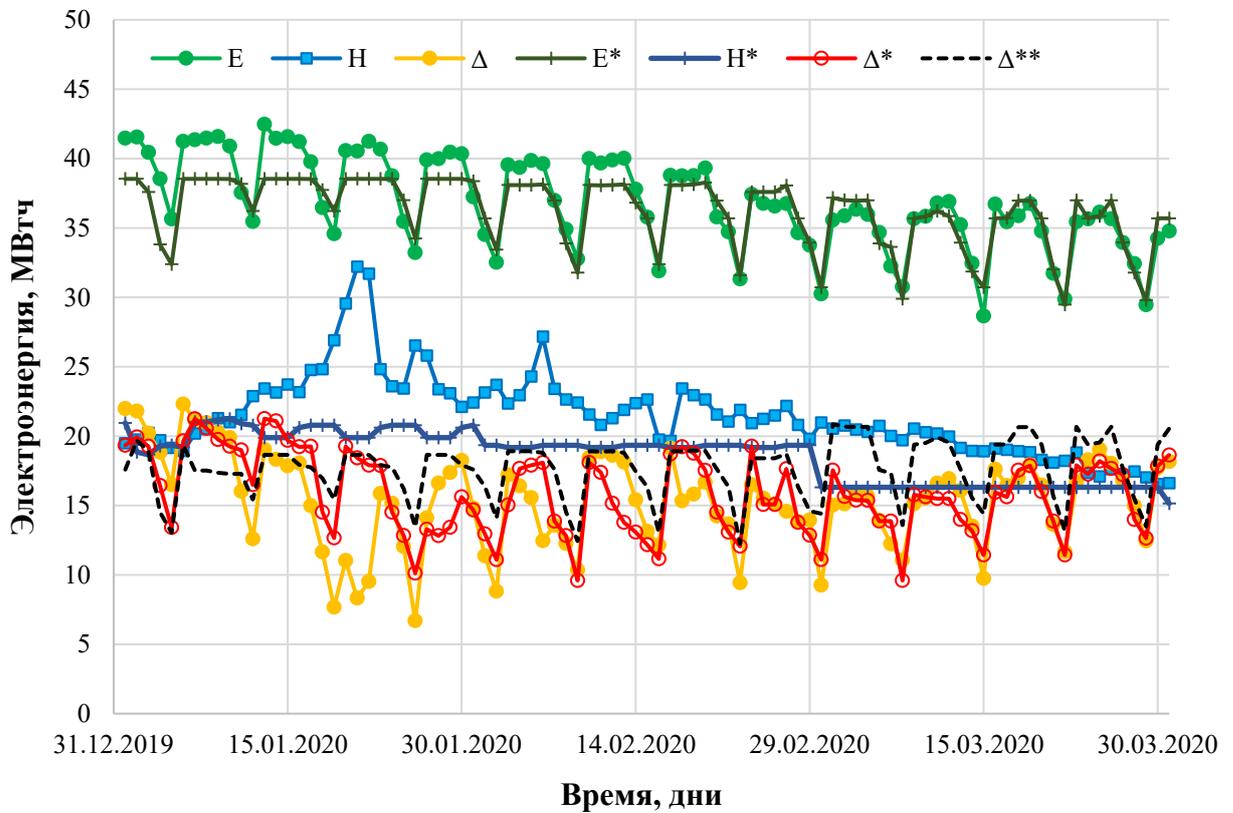


Рисунок 4.10 – Результаты прогноза AdaBoost, январь – март 2020 г.

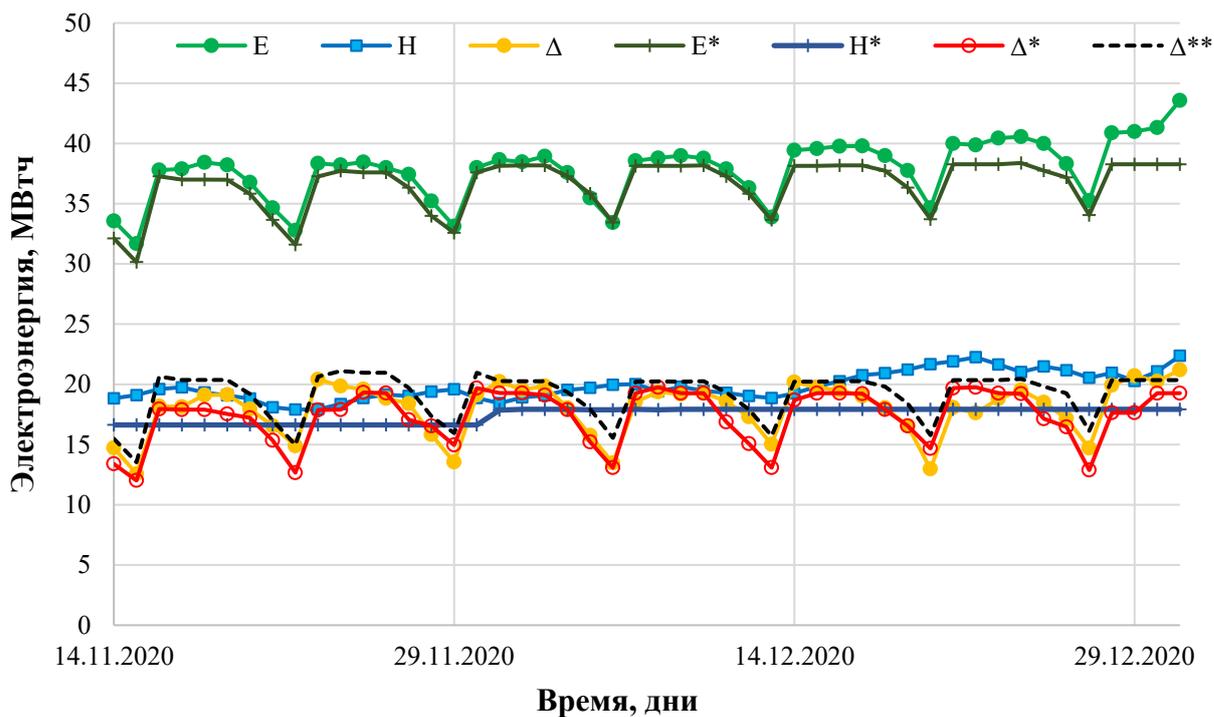


Рисунок 4.11 – Результаты прогноза AdaBoost, январь – март 2020 г.

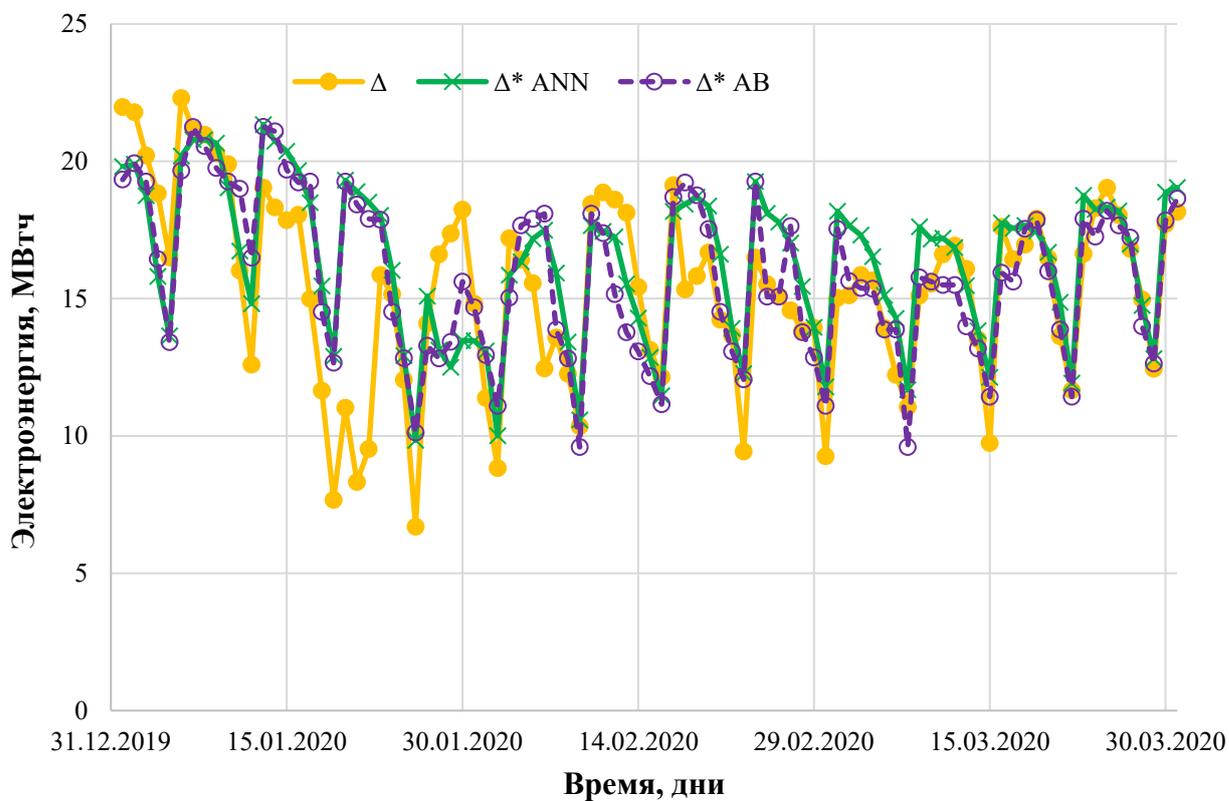


Рисунок 4.12 – Сопоставление прогнозов ANN и AdaBoost и истинным значением разности потребления и генерации ГЭС

У нейросетевой модели больше отклонений в сторону завышения прогноза, что ведет к излишнему использованию ДЭС, что показано на Рисунке 4.13, где приведены суммарные значения запланированной выработки ДЭС за каждые две недели (суммирование за две недели выбрано для большей наглядности). В то же время ансамблевая модель не приводит к риску чрезмерного увеличения выработки ГЭС, так как ее ошибки в большую и меньшую сторону распределены равномерно как суммарно, так и во времени. Это можно объяснить тем, что модель на базе деревьев решений из-за своей дискретной природы принципиально строит прогноз, сглаживающий как пики, так и провалы.

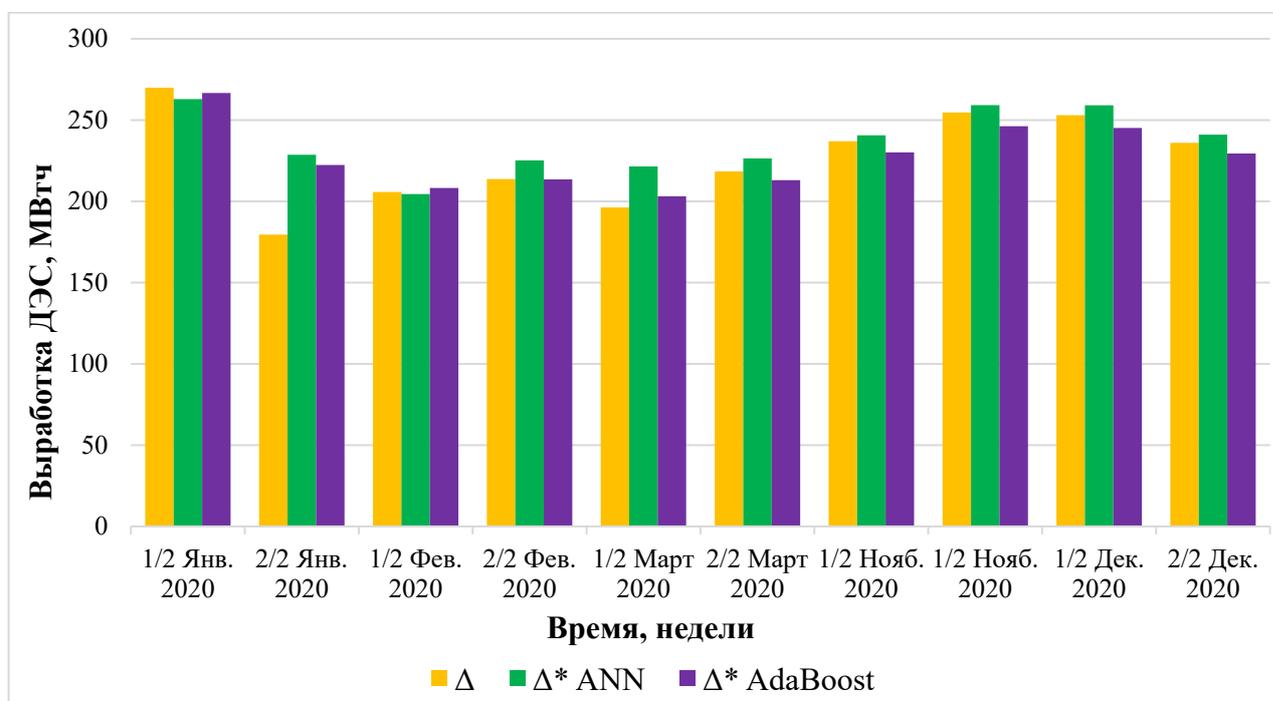


Рисунок 4.13 – Суммарные объемы генерации ДЭС в соответствии с планами, полученными от ANN и AdaBoost (шаг времени для суммирования – две недели)

Таким образом показано, что выбор модели на основании только ее средней ошибки прогнозирования может привести к неэффективному использованию ДЭС. На тестовом интервале избыточная выработка ДЭС нейросетевой моделью по сравнению с ансамблевой моделью AdaBoost составила 63 МВтч за рассматриваемый период, что составляет 453 кВтч в сутки или 18,89 кВт, если рассматривать не электроэнергию, а мощность. При этом рассмотрены лучшие

результаты обеих моделей. Если рассмотреть сценарий использования AdaBoost при выборе неудачной для него стратегии прогнозирования не непосредственно разности потребления и генерации ГЭС, а использование отдельных прогнозов потребления и генерации, то разность $S_{\delta+}$ составит 166 МВтч (281 – 115), что составляет 1194 кВтч в сутки или 49,76 кВт.

При этом точность планирования использования ДЭС можно повысить, если использовать данные о притоке в водохранилище [135] и уровне воды в нем. Это позволит достичь еще большего снижения использования ДЭС по сравнению с используемыми сейчас менее точными способами планирования.

4.5 Расчет экономических и экологические показатели

Чтобы оценить качество полученных планов генерации, следует определить экономический и экологический эффекты от использования каждого конкретного плана. Поскольку цель расчета заключается в сопоставлении найденных альтернативных решений, из анализа можно исключить постоянные издержки на выработку электроэнергии. Также, с учетом отмеченных в первой части работы высоких региональных цен на дизельное топливо, допустимо не рассматривать расходы ГЭС на генерацию мощности. Дополнительным аргументом в данном случае служит и то, что эксплуатация уже построенной ГЭС, как возобновляемого и чистого источника энергии, никак не влияет на экологическую обстановку в ГБАО. На этом основании планы далее оцениваются по затратам топлива ДЭС и выбросам этого типа генераторов в атмосферу на рассматриваемом интервале времени.

Прежде чем выполнить намеченный расчет затрат необходимо задаться недельными зимним и осенним почасовыми графиками потребления для ГБАО. Ретроспективные траектории активной мощности нагрузки ГБАО в относительных единицах для зимнего и весенне-осеннего периода представлены соответственно на рисунках 4.14 и 4.15.

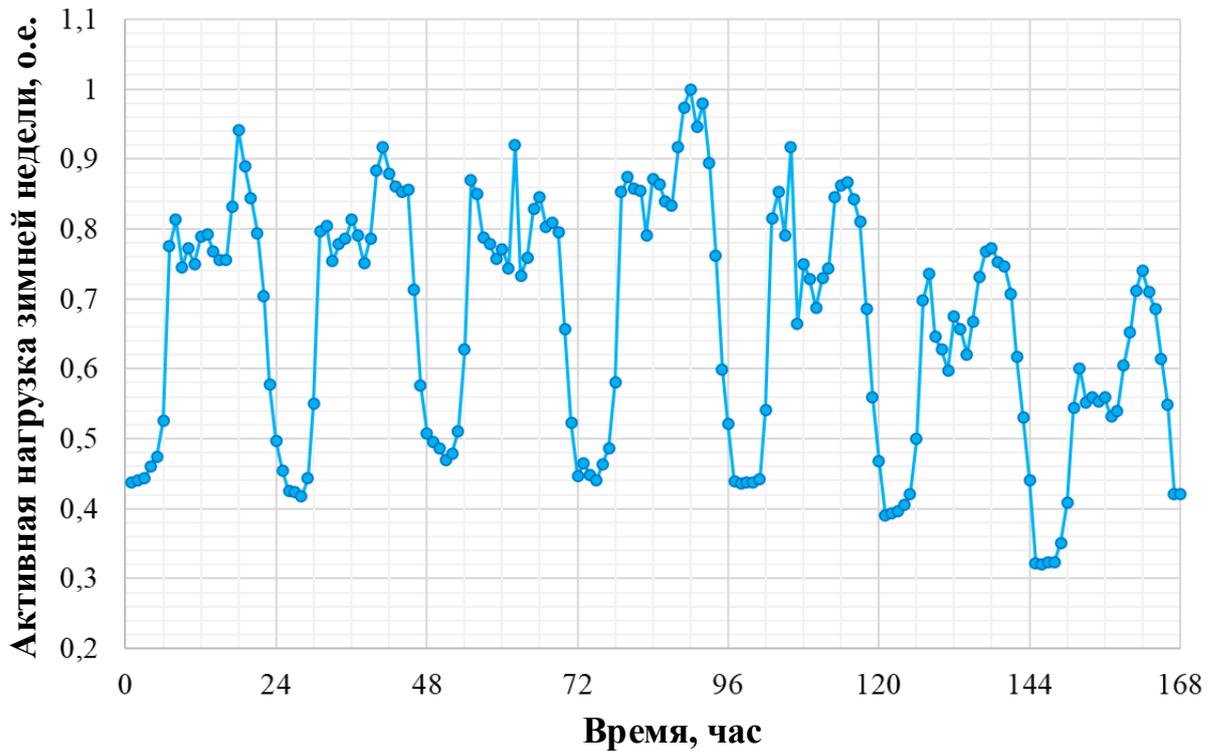


Рисунок 4.14 – Недельный график электропотребления для недели зимнего сезона

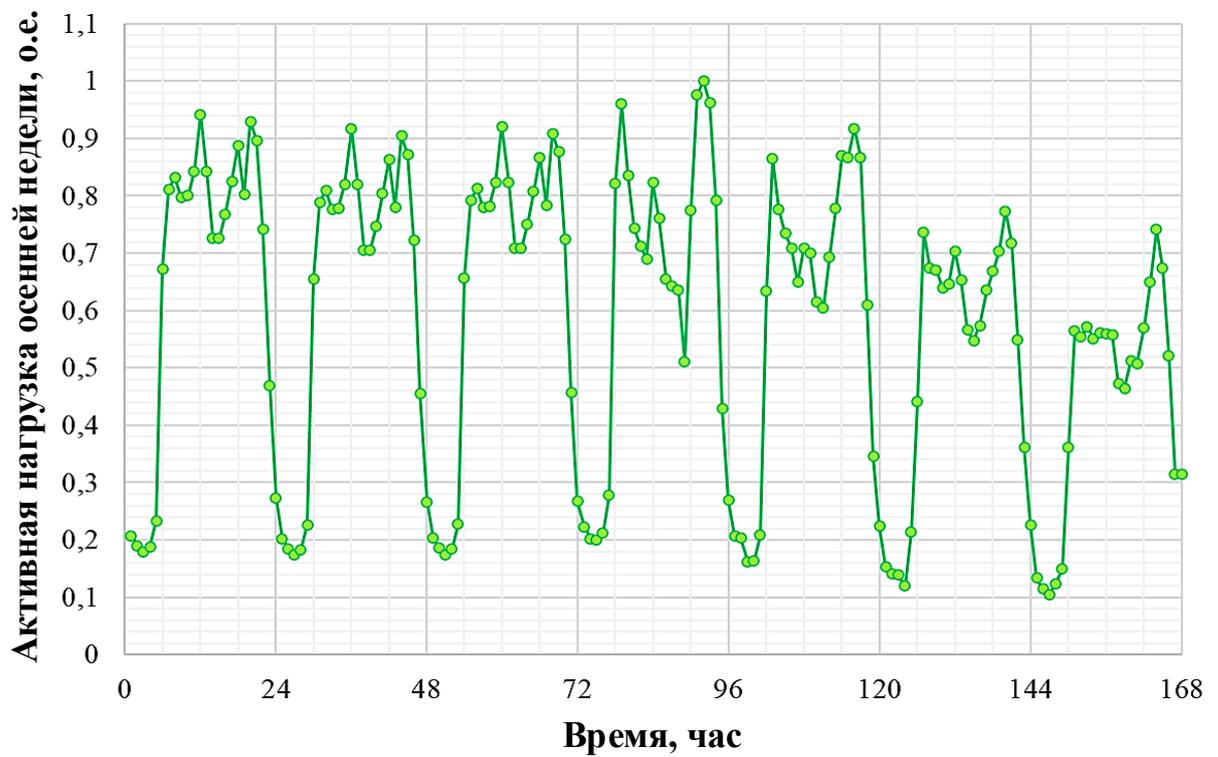


Рисунок 4.15 – Недельный график электропотребления для недели весенне-осеннего периода

По представленным графикам определяется характерное число часов использования максимума нагрузки для каждого дня недели T_{max}^d . Оценки параметра представлены в таблице 4.6.

Таблица 4.6 – Характерное число часов использования максимума нагрузки по дням недели

	Сезон	Пн	Вт	Ср	Чт	Пт	Сб	Вс
T_{max}^d , час	Весна/Осень	16.54	16.54	16.54	15.12	15.38	15.85	14.37
	Зима	17.60	18.35	18.26	18.08	17.84	18.76	17.00

Для оценки расхода дизельного топлива $B_{g,t}$ генератором g за час t используется эмпирическое выражение, предложенное и использованное в работах [136]:

$$B_{g,t} = A_g P_{g,t} + B_g P_{g.rated}, \quad (4.24)$$

где $P_{g,t}$ и $P_{g.rated}$ – режимная в час t и установленная мощности генератора g , кВт; $A_g = 0,246$ л/кВт и $B_g = 0,08145$ л/кВт – коэффициенты модели расхода топлива. Как можно видеть, затраты на топливо зависят не только от загрузки в конкретном режиме, но и от состава включенного оборудования. Если предположить, что состав генераторов для рассматриваемых планов одинаков, то вторую составляющую уравнения (4.24) можно отбросить, поскольку в таких условиях она, очевидно, будет одинакова для всех вариантов и не повлияет на результаты сопоставления и фактически входит в постоянные издержки. При этом такое решение позволяет без значимых допущений отказаться от промежуточного распределения нагрузки между ДЭС ГБАО и решения задачи выбора состава оборудования даже в упрощенном виде. Тогда совокупные переменные затраты за исследуемый период рассчитываются следующим образом:

$$B_{\Sigma} = \sum_{t=1}^T B_{g,t} = \sum_{d=1}^D A_g P_d^{max} T_{max,d}, \quad (4.25)$$

где P_d^{max} – максимальная вырабатываемая ДЭС мощность в день d , кВт;
 $T_{max,d}$ – число часов максимума нагрузки в день d , час.

Расчет финансовых затрат производится исходя из фиксированной цены на дизельное топливо для ГБАО – $c_0^F = 90$ руб/л. Параметр c_0^F учитывает в том числе и транспортные затраты на доставку топлива. Итоговые переменные затраты в таком случае можно определить исходя из B_Σ :

$$C_\Sigma = c_0^F B_\Sigma. \quad (4.26)$$

Для расчета объема выбрасываемых в атмосферу парниковых газов при использовании ДЭС используются статистические данные по загрязнению, систематизированные в исследованиях [137-140]. Так, в частности, указано, что удельные выбросы на сжигание литра дизельного топлива с учетом углеродного следа на его производство составляют $m_0^{GHG} = 3,15$ kg/l. Расчет переменной составляющей объема выбросов при известном расходе топлива выполняется просто:

$$M_\Sigma = m_0^{GHG} B_\Sigma. \quad (4.27)$$

По формулам (4.24)–(4.27) для трех планов рассчитаны расход топлива и затраты на него, а также суммарные выбросы парниковых газов от работы ДЭС. Результаты расчета для сформированных графиков генерации представлены в таблице 4.7.

Таблица 4.7 – Результаты оценки степени экологичности и экономичности найденных планов генерации

	B_Σ, l	Δ_B, l	C_Σ , тыс. руб	Δ_C , тыс. руб	M_Σ , kg	Δ_{GHG} , kg	Δ , %
Δ^* АВ	560 328	–	50429,52	–	1 118 195	–	–
Δ^* ANN	582 822	22 494	52454,04	2024,46	1 138 097	19 903	4.01
Δ^{**}	587 495	27 167	52874,58	2445	1 153 216	35 021	4.85

На рисунке 4.16 представлены предполагаемые графики выбросов парниковых газов для трех рассматриваемых сценариев. По выражениям (4.24)–(4.27) и результатам в Таблице 4.7 можно заключить, что объем выбросов, расход

топлива и затраты пропорциональны, поэтому суточный график для последних величин будет иметь ту же форму, что графики на рисунке ниже.

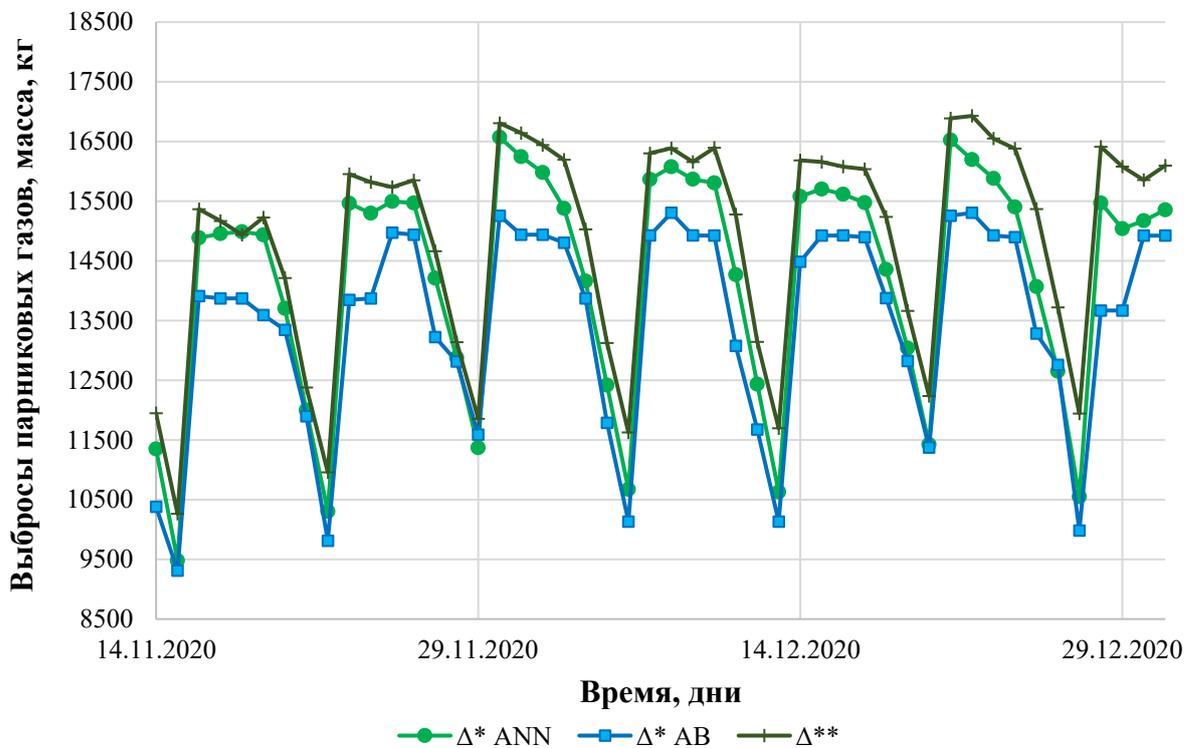


Рисунок 4.16 – Графики выбросов ДЭС ГБАО для трех рассматриваемых сценариев

Выводы по главе 4

Средняя ошибка среднесрочного прогнозирования разности потребления энергосистемы и генерации ГЭС в холодный сезон для рассматриваемой изолированной энергосистемы составила 1,5 МВтч или 9 %. Определено, что метеорологический прогноз на неделю вперед не играет существенной роли в прогнозировании как потребления, так и выработки ГЭС. Поэтому для рассматриваемого объекта необходимо совершенствовать сеть метеорологических постов и методы прогнозирования температуры, скорости ветра, осадков. Точность прогноза оказывает существенное влияние на эффективность использования дизельного топлива ДЭС. Завышение прогноза разности потребления и генерации приводят к планированию избыточной выработки ДЭС и холостым сбросам ГЭС.

В то же время введение мер, занижающих такой прогноз, может привести к дефициту электроэнергии.

Показано, что в зависимости от выбора модели и стратегии планирования генерации ДЭС (использование прогнозов потребления и выработки ГЭС для нахождения их разности или непосредственное прогнозирование их разности) излишняя выработка ДЭС за холодный сезон (ноябрь-март) может составлять от 100 до 300 МВтч (720 – 2160 кВтч за сутки). Нужно отметить, что в работе принят ряд упрощений, в частности, использован график реальной выработки ГЭС в качестве опорного, хотя он не является оптимальным. Кроме того, ограничения на регулирование генерации ГЭС рассмотрены упрощенно. Поэтому полученные результаты по избыточному использованию ДЭС на данный момент являются оценочными. Тем не менее полученные результаты обосновывают реально существующую проблему перерасхода дизельного топлива в изолированной энергосистеме из-за погрешностей планирования баланса генерации и потребления, а также показывают перспективы снижения выбросов CO₂ и расходов на ископаемое топливо за счет повышения точности прогнозирования и оптимизации регулирования выработкой ГЭС.

ЗАКЛЮЧЕНИЕ

Основные результаты работы заключаются в следующем:

1. Показано, что регион богат возобновляемыми энергетическими ресурсами, особенно гидроэнергетикой. Широкое использование солнечной энергии оказало бы положительное влияние на энергетическую и экономическую ситуацию в регионе. Что касается энергии ветра, то в настоящее время она экономически неконкурентоспособна с действующей гидроэнергетикой и солнечной генерацией. Необходимость увеличения доли микро-, мини- и малых ГЭС и сокращения доли существующих ДЭС в зимний период, поскольку в регионе практически нет месторождений угля, нефти и газа.

2. Установлено, что для задач среднесрочного прогнозирования электропотребления и генерации, целесообразно использовать методы, основанные на машинном обучении.

3. Доказана возможность повышения точности прогнозирования электропотребления обеспечивается использованием метеорологических данных и кластеризации метеорологических условий. Разработана новая эффективная модель среднесрочного прогнозирования электропотребления. На этапе предобработки данных применен метод разложения пространства признаков на главные компоненты, который, во-первых, позволяет проверить корректность кластеризации, во-вторых, агрегировать метеорологические признаки. В результате средняя ошибка прогноза на тестовой части выборки данных составила 0.59 МВт, среднеквадратическая ошибка 0.77 МВт, средняя ошибка в процентах 4.5 %. По сравнению с авто регрессионной моделью без кластеризации и учета метеорологических факторов (MAE 0.80 МВт, RMSE 0.96 МВт, MAPE 6.11 %).

4. Предложено и обосновано использование компактных моделей на основе ансамблей неглубоких деревьев решений для решения задачи среднесрочного прогнозирования притока воды в условиях дефицита достоверной информации в работе. Разработан комплекс авторских моделей среднесрочного прогнозирования притока воды в водохранилища ГЭС на основе ансамблевых методов машинного

обучения для определения объемов воды, достаточных для регулирования генерации в изолированной системе. Ошибка (RMSE) ансамблевых моделей за счет самоадаптации сократилась с 4.4–4.5 м³/с до 4.0 м³/с и оказалась ниже ошибки более сложной многослойной рекуррентной модели (4.9 м³/с). В самый трудный для прогнозирования год за счет непрерывной самоадаптации ошибка ансамблевых моделей сократилась с 8.3–8.6 м³/с до 5.3–6.3 м³/с.

5. Для решения задачи обеспечения балансовой надежности изолированных энергосистем с высокой долей малых ГЭС деривационного типа выполнен анализ семи различных моделей для прогнозирования генерации гидроэлектростанциями на основе ретроспективных данных о генерации и метеорологических параметров. Предложена новая модель комплексного планирования графика генерации на среднесрочную перспективу на основе адаптивного бустинга линейных регрессий с регуляризацией Тихонова, обеспечивающая близкий к оптимальному выбор состава генерирующего оборудования для обеспечения балансовой надежности. Средняя ошибка на тестовой выборке составила 5,23 % или 44 МВт.ч при средней суточной генерации 718 МВт.ч.

6. Определено, что метеорологический прогноз на неделю вперед не играет существенной роли в прогнозировании как потребления, так и выработки ГЭС. Поэтому для рассматриваемого объекта необходимо совершенствовать сеть метеорологических постов и методы прогнозирования температуры, скорости ветра, осадков. Точность прогноза оказывает существенное влияние на эффективность использования дизельного топлива ДЭС. Завышение прогноза разности потребления и генерации приводят к планированию избыточной выработки ДЭС и холостым сбросам ГЭС. В то же время введение мер, занижающих такой прогноз, может привести к дефициту электроэнергии.

7. Показано, что в зависимости от выбора модели и стратегии планирования генерации ДЭС (использование прогнозов потребления и выработки ГЭС для нахождения их разности или непосредственное прогнозирование их разности) излишняя выработка ДЭС за холодный сезон (ноябрь-март) может составлять от 100 до 300 МВтч (720 – 2160 кВтч за сутки). Нужно отметить, что в работе принят

ряд упрощений, в частности, использован график реальной выработки ГЭС в качестве опорного, хотя он не является оптимальным. Кроме того, ограничения на регулирование генерации ГЭС рассмотрены упрощенно. Поэтому полученные результаты по избыточному использованию ДЭС на данный момент являются оценочными. Тем не менее полученные результаты обосновывают реально существующую проблему перерасхода дизельного топлива в изолированной энергосистеме из-за погрешностей планирования баланса генерации и потребления, а также показывают перспективы снижения выбросов CO_2 и расходов на ископаемое топливо за счет повышения точности прогнозирования и оптимизации регулирования выработкой ГЭС.

Перспективы дальнейших исследований заключаются в решении задачи планирования ДЭС и регулирования генерации ГЭС с учётом накопителей энергии в виде оптимального динамического управления в условиях неопределенности.

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

ИЭС – изолированная энергосистема

ГБАО – Горно-Бадахшанская автономная область

ГЭС – гидроэлектростанция

ДЭС – дизельная электростанция

ВИЭ – возобновляемый источник энергии

ВЭС – ветровая электростанция

СЭС – солнечная электростанция

LR – линейная регрессия

ANN – искусственная нейронная сеть

RF – случайный лес

AB – бустинг

ABLR – адаптивный бустинг линейных моделей

XGB – экстремальный градиентный бустинг

SVR – регрессия опорного вектора

SVM – метод опорных векторов

kNN – k-ближайших соседей

MLP – многослойный персептрон

LSTM – блок долгой краткосрочной памяти

СПИСОК ЛИТЕРАТУРЫ

1. Science and policy characteristics of the Paris Agreement temperature goal / C.F. Schleussner, J. Rogelj, M. Schaeffer [et al.] // *Nature Climate Change*. – 2016. – Vol. 6, iss. 9. – P. 827-835.
2. Paris Agreement climate proposals need a boost to keep warming well below 2⁰C / J. Rogelj, M. den Elzen, N. Höhne [et al.] // *Nature*. – 2016. – Vol. 534, iss. 7609. – P. 631-639.
3. Nikolic, D. Adding inertia to isolated power systems for 100% renewable operation / D. Nikolic, M. Negnevitsky // *Energy Procedia*. – 2019. – Vol. 159. – P. 460-465.
4. Panwar, N.L. Role of renewable energy sources in environmental protection: A review / N.L. Panwar, S.C. Kaushik, S. Kothari // *Renewable and sustainable energy reviews*. – 2011. – Vol. 15, iss. 3. – P. 1513-1524.
5. Akella, A.K. Social, economical and environmental impacts of renewable energy systems / A.K. Akella, R.P. Saini, M.P. Sharma // *Renewable Energy*. – 2009. – Vol. 34, iss. 2. – P. 390-396.
6. Solaun, K. Climate change impacts on renewable energy generation. A review of quantitative projections / K. Solaun, E. Cerdá // *Renewable and sustainable energy Reviews*. – 2019. – Vol. 116. – P. 109415.
7. Разработка моделей среднесрочного прогнозирования электропотребления в изолированно работающих энергосистемах на основе ансамблевых методов машинного обучения / С.М. Асанова, Дж.С. Ахъёев, С.А. Дмитриев и др. // *Известия НТЦ Единой энергетической системы*. – 2021. – № 1. – С. 32-39.
8. Разработка моделей среднесрочного прогнозирования притока воды для планирования выработки электроэнергии в изолированных энергосистемах / М.Х. Сафаралиев, С.Е. Кокин, С.А. Дмитриев и др. // *Электричество*. – 2022. – № 5. – С. 58-68.
9. Адаптивные ансамблевые модели для среднесрочного прогнозирования выработки электроэнергии гидроэлектростанциями в изолированных энергосистемах с учётом изменений температуры / М.Х. Сафаралиев, П.В.

- Матренин, С.А. Дмитриев и др. // Электротехнические системы и комплексы. – 2022. – № 1 (54). – С. 38-45.
10. Ашуров, Д. Использование альтернативных источников питания горных населенных пунктов Республики Таджикистан / Д. Ашуров // Инновации в технологиях и образовании. – 2016. – С. 142-144.
 11. A comparative life cycle assessment of electric, compressed natural gas, and diesel buses in Thailand / N.R. Gabriel, K.K. Martin, S.J. Haslam [et al.] // Journal of Cleaner Production. – 2021. – Vol. 314. – P. 128013.
 12. Kapustin, A. Results of assessing CO₂ emissions from e-vehicles in case of their possible switching to electricity / A. Kapustin, V. Rakov // Transportation research procedia. – 2018. – Vol. 36. – P. 266-273.
 13. Techno-economic assessment of photovoltaic-diesel generator-battery energy system for base transceiver stations loads in Nigeria / O.M. Babatunde, I.H. Denwigwe, D.E. Babatunde [et al.] // Cogent Engineering. – 2019. – Vol. 6, iss. 1. – P. 1684805.
 14. Гуломзода, А.Х. Модифицированный способ синхронизации Microgrid с внешней изолированной энергосистемой / А.Х. Гуломзода, М.Х. Сафаралиев, Е.А. Люханов // Электротехнические системы и комплексы. – 2021. – № 3 (52). – С. 72–80.
 15. Сангов, Х.С. Применение солнечных фотоэлектрических установок в системах электроснабжения сельских потребителей (На примере удаленных регионов Республики Таджикистан) / Х.С. Сангов // Политехнический вестник. Серия: Инженерные исследования. – 2020. – №4. – С. 30-38.
 16. Солнечные панели на «Крыше мира» [Электронный ресурс]. – Режим доступа: <https://www.usaid.gov/ru/tajikistan/success-stories/dec-2020-installing-solar-panels-world-rooftop>.
 17. Киргизов, А.К. Развитие и оптимизация режимов электроэнергетической системы с распределенными возобновляемыми источниками энергии методами искусственного интеллекта (на примере Республики Таджикистан): дис. ... канд. техн. наук: 05.14.02 / Киргизов Алифбек Кирнизович. – Томск,

2017. – 189 с.
18. Закон Республики Таджикистан от 12 января 2010 года № 587 «Об использовании возобновляемых источников энергии» [Электронный ресурс]. – Режим доступа: https://base.spinform.ru/show_doc.fwx?rgn=30439.
 19. Постановление Совета Министров Республики Таджикистан от 5 апреля 1993 года № 139 «О мерах по стимулированию развития малой энергетики и увеличения добыча угля в Республике Таджикистан» [Электронный ресурс]. – Режим доступа: http://www.adlia.tj/show_doc.fwx?Rgn=9962.
 20. Постановление Правительством Республики Таджикистан от 4 июня 1997 года № 267 «О развитии малой энергетики Республики Таджикистан» [Электронный ресурс]. – Режим доступа: http://www.adlia.tj/show_doc.fwx?Rgn=6452.
 21. Худжасаидов, Дж.Х. Анализ и планирование режимов электроэнергетической системы с каскадом гидроэлектростанций (на примере электроэнергетической системы Памира): дис. ... канд. техн. наук: 05.14.02 / Худжасаидов Джахонгир Худжасаидович. – Новосибирск, 2018. – 183 с.
 22. Петров, Г.Н. Малая Гидроэнергетика Таджикистана / Г.Н. Петров, Х.М. Ахмедов // Гидротехническое строительство. – 2010. – № 12. – С. 49-56.
 23. Electromagnetic transients in the control system of output parameters of a solar power plant in Tajikistan Central Asia region / B.N. Sharifov, M.Kh Safaraliev, V.Z. Manusov [et al.] // International Journal of Hydrogen Energy. – 2022. – Vol. 47, iss. 9. – P. 5757-5765.
 24. Оценка энергетического потенциала солнечной радиации региона с применением солнечного трекера / В.З. Манусов, Дж.С. Ахъеев, М.Х. Назаров и др. // Научный вестник Новосибирского государственного технического университета. – 2020. – № 1. – С. 189-203.
 25. Energy potential estimation of the region's solar radiation using a solar tracker / M.K. Safaraliev, I.N. Odinaev, J.S. Ahyoev [et al.] // Applied Solar Energy. – 2020. – Vol. 56. – iss. 4. – P. 270-275.
 26. Laldjebaev, M. Renewable energy in Central Asia: An overview of potentials,

- deployment, outlook, and barriers / M. Laldjebaev, R. Isaev, A. Saukhimov // *Energy Reports*. – 2021. – Vol. 7. – P. 3125-3136.
27. Sangov, Kh.S. Integration of Renewable Energy Sources into the Power Supply System of the Murghab Settlement, Gorno-Badakhshan Autonomous Region / Kh.S. Sangov, S.N. Chorshanbiev // 2022 Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus): proc., Saint Petersburg, 25-28 January 2022. – Saint Petersburg: IEEE, 2022. – P. 1260-1263.
28. Бартоломей, П.И. Оптимизация режимов энергосистем / П.И. Бартоломей, Т.Ю. Паниковская. – Екатеринбург: УГТУ–УПИ, 2008. – 164 с.
29. Hahn, H. Electric load forecasting methods: Tools for decision making / H. Hahn, S. Meyer-Nieberg, S. Pickl // *European journal of operational research*. – 2009. – Vol. 199, iss. 3. – P. 902-907.
30. Piotrowski, P. Forecasting in Power Engineering in different time horizons / P. Piotrowski // *Prace Naukowe PW “Elektryka”*. – 2013. – № 144.
31. Very short-term load forecasting: wavelet neural networks with data pre-filtering / C. Guan, P.B. Luh, L.D. Michel [et al.] // *IEEE Transactions on Power Systems*. – 2012. – Vol. 28, iss. 1. – P. 30-41.
32. Jiang, H. Short-term load forecasting based automatic distribution network reconfiguration / H. Jiang, F. Ding, Y. Zhang // 2017 IEEE Power & Energy Society General Meeting: proc., Chicago, 16-20 July 2017. – Chicago: IEEE, 2017. – P. 1 - 5.
33. İlseven, E. Medium-term electricity demand forecasting based on MARS / E. İlseven, M. Göl // 2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe): proc., Turin, 26-29 September 2017. Turin: IEEE, 2017. – P. 1-6.
34. Long term load forecasting accuracy in electric utility integrated resource planning / J.P. Carvallo, P.H. Larsen, A.H. Sanstad, Ch.A. Goldman // *Energy Policy*. – 2018. – Vol. 119. – P. 410-422.
35. Electricity load forecasting: A systematic review / I.K. Nti, M. Teimeh, Nyarko-Boateng, A.F. Adekoya // *Journal of Electrical Systems and Information*

- Technology. – 2020. – Vol. 7, iss. 1. – P. 1-19.
36. Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series / M. Espinoza, C. Joye, R. Belmans, B. De Moor // IEEE Transactions on Power Systems. – 2005. – Vol. 20, iss. 3. – P. 1622-1630.
 37. Blind kalman filtering for short-term power load forecasting / S. Sharma, A. Majumdar, V. Elvira, É. Chouzenoux // IEEE Transactions on Power Systems. – 2020. – Vol. 35, iss. 6. – P. 4916-4919.
 38. Arora, S. Short-term forecasting of anomalous load using rule-based triple seasonal methods / S. Arora, J.W. Taylor // IEEE Transactions on Power Systems. – 2013. – Vol. 28, iss. 3. – P. 3235-3242.
 39. Multi-scale convolutional neural network with time-cognition for multi-step short-term power load forecasting / Z. Deng, B. Wang, Y. Xu [et al.] // IEEE Access. – 2019. – Vol. 7. – P. 88058-88071.
 40. Chen, B.-J. Load forecasting using support vector machines: a study on EUNITE competition 2001 / B.-J. Chen, M.-W. Chang, Ch.-J. Lin // IEEE Transactions on Power Systems. – 2004. – Vol. 19, iss. 4. – P. 1821-1830.
 41. A short-term power load forecasting method using integrated CNN and LSTM network / S.H. Rafi, Nahid-Al-Masood, S.R. Deeba, E. Hossain // IEEE Access. – 2021. – Vol. 9. – P. 32436-32448.
 42. Sergeev, N.N. Enhancing Efficiency of Ensemble Machine Learning Models for Short-Term Load Forecasting through Feature Selection / N.N. Sergeev, P.V. Matrenin // 2022 IEEE 23rd International Conference of Young Professionals in Electron Devices and Materials (EDM): proc., Altai, 30 June 2022 - 04 July 2022. – Altai: IEEE, 2022. – P. 368-371.
 43. Draper, N.R. Applied regression analysis / N.R. Draper, H. Smith. – Wiley: John Wiley & Sons, 1998.
 44. Weisberg, S. Applied linear regression / S. Weisberg. – Wiley: John Wiley & Sons, 2005. – 368 p.
 45. Heineman, G.T. The Relationship Between Summer Weather and Summer Loads –

- A Regression Analysis / G.T. Heineman, D.A. Nordman, E.C. Plant // IEEE Transaction Power Apparatus System. – 1966. – Vol. PAS-85, iss. 11. – P. 1144-1154.
46. Al-Hamadi, H.M. Long-term/mid-term electric load forecasting based on short-term correlation and annual growth / H.M. Al-Hamadi, S.A. Soliman // Electr. Power Syst. Res. – 2005. – Vol. 74, iss. 3. – P. 353-361.
 47. Nonlinear regression modeling of nutrient loads in streams: A Bayesian approach / S.S. Qian, K.H. Reckhow, J. Zhai, G. McMahon // Water Resources Research. – 2005. – Vol. 41, iss. 7. – Art. W07012.
 48. Raza, M. On recent advances in PV output power forecast / M. Raza, M. Nadarajah, Ch. Ekanayake // Solar Energy. – 2016. – Vol. 136. – P. 125-144.
 49. Mbamalu, G.A.N. Load Forecasting Via Suboptimal Seasonal Autoregressive Models and Iteratively Reweighted Least Squares Estimation / G.A.N. Mbamalu, M.E. El-Hawary // IEEE Transaction on Power System. – 1992. – Vol. 8. – P. 343-348.
 50. Estimating Substation Peaks from Research Data / R.P. Broadwater, A. Sargent, A. Yarali [et al.] // IEEE Transaction on Power Delivery. – 1997. – Vol. 12. – P. 451-456.
 51. El-Keib, A.A. Advancement of Statistical Based Modeling for Short-Term Load Forecasting / A.A. El-Keib, X. Ma, H. Ma // Electric Power Systems Research. – 1995. – Vol. 35. – P. 51-58.
 52. Infield, D.G. Optimal Smoothing for Trend Removal in Short term Electricity Demand Forecasting / D.G. Infield, D.C. Hill // IEEE Transaction on Power System. – 1998. – Vol. 13. – P. 1115-1120.
 53. Probability concepts explained: Maximum likelihood estimation [Electronic resource]. – Available from: <https://towardsdatascience.com/probability-concepts-explained-maximum-likelihood-estimation-c7b4342fdbb1>.
 54. Babu, S.K. (2018). Maximum Likelihood Robust Regression Model (MLRRM) for Downscaling data sets in Hydrology / S.K. Babu, M. Srinivasulu, M.V. Ramanaiah // International Journal of Statistics and Systems. – 2018. – Vol. 13, iss. 2. – P. 95-

- 100.
55. A hybrid statistical method to predict wind speed and wind power / H. Liu, H.-Q. Tian, C. Chen, Y.-f. Li // *Renew Energy*. – 2010. – Vol. 35, iss. 8. – P. 1857-1861.
 56. Laurent, P. Bayesian statistical analysis applied to solar radiation modelling / P. Laurent, J. Boland, B. Ridley // *Renew Energy*. – 2013. – Vol. 49. – P. 124-127.
 57. Short term wind speed predictions by using the grey prediction model based forecast method / C.-Y. Huang, Y.-W. Liu, W.-C. Tzeng, P.-Y. Wang // 2011 IEEE Green Technologies Conference (IEEE-Green): proc., Baton Rouge, 14-15 April 2011. – Baton Rouge: IEEE, 2011. – P. 1-5.
 58. Bracale, A. An advanced bayesian method for short-term probabilistic forecasting of the generation of wind power / A. Bracale, P. De Falco // *Energies*. – 2015. – Vol. 8, iss. 9. – P. 10293-10314.
 59. El-Fouly, T.H. One day ahead prediction of wind speed and direction / T.H. El-Fouly, E.F. El-Saadany, M.M. Salama // *IEEE Trans Energy Convers*. – 2008. – Vol. 23, iss. 1. – P. 191-201.
 60. Wang, M.-D. Short-term wind speed forecasting combined time series method and arch model / M.-D.Wang, Q.-R. Qiu, B.-W. Cui // 2012 international conference on machine learning and cybernetics (ICMLC): proc., Xi'an, 15-17 July 2012. – Xi'an: IEEE, 2012. – P. 924-927
 61. Wind speed forecasting based on second order blind identification and autoregressive model / U. Firat, S.N. Engin, M. Saraclar, A.B. Ertuzun // 2010 Ninth International Conference on Machine Learning and Applications (ICMLA): proc., Washington, 12-14 December 2010. – Washington: IEEE, 2010. – P. 686-691.
 62. The use of wavelet theory and arma model in wind speed prediction / Li Ling-ling, J.-H. Li, P.-J. He, C.-S. Wang // 2011 1st International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST): proc., Xi'an, 23-27 October 2011. – Xi'an: IEEE, 2011. – P. 395-398.
 63. Azad, H.B. Long-term wind speed forecasting and general pattern recognition using neural networks / H.B. Azad, S. Mekhilef, V.G. Ganapathy // *IEEE Trans Sustain Energy*. – 2014. – Vol. 5, iss. 2. – P. 546-553.

64. Vapnik, V.N. *Statistical Learning Theory* / V.N. Vapnik. – New York: Wiley, 1998.
65. Chen, B.J. *Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001* / B.J. Chen, M.W. Chang, C.J. Lin // *IEEE Trans. Power Syst.* – 2004. – Vol. 19, iss. 4. – P. 1821-1830.
66. Tay, F.E.H. *Modified Support Vector Machines in Financial Time Series Forecasting* / F.E.H. Tay, L.J. Cao // *Neurocomputing.* – 2002. – Vol. 48. – P. 847-861.
67. Zhu, J. *Using markov chains for link prediction in adaptive web sites* / J. Zhu, J. Hong, J.G. Hughes // *Int. Conference on Soft Issues in the Design, Development, and Operation of Computing Systems: proc., Berlin, Heidelberg, 10 April 2002.* – Berlin, Heidelberg: Springer, 2002. – P. 60-73.
68. Cover, T. *Nearest neighbor pattern classification* / T. Cover, P. Hart // *IEEE Trans. Inf. Theory.* – 1967. – Vol. 13. – P. 21-27.
69. *Potential of Artificial Neural Network to Power System Operation* / M.J. Damborg, M.A El-Sharkawi, M.E. Aggoune, R.J. Marks II // *IEEE International Symposium on Circuits and Systems: proc., New Orleans, LA, 01-03 May 1990.* – New Orleans, LA: IEEE, 1990. – P. 2933-2937.
70. Zhang, G. *Forecasting with artificial neural networks: The state of the art* / G. Zhang, B.E. Patuwo, M.Y. Hu // *International journal of forecasting.* – 1998. – Vol. 14, iss. 1. – P. 35-62.
71. Inman, R.H. *Solar forecasting methods for renewable energy integration* / R.H. Inman, H.T. Pedro, C.F. Coimbra // *Prog Energy Combust Sci.* – 2013. – Vol. 39, iss. 6. – P. 535-576.
72. Doucoure, B. *Time series prediction using artificial wavelet neural network and multi-resolution analysis application to wind speed data* / B. Doucoure, K. Agbossou, A. Cardenas // *Renew Energy.* – 2016. – Vol. 92. – P. 202-211.
73. *A hybrid neuro-fuzzy power prediction system for wind energy generation* / A.E. Saleh, M.S. Moustafa, K.M. Abo-Al-Ez, A.A. Abdullah // *Int J Electr Power Energy Syst.* – 2016. – Vol. 74. – P. 384-395.
74. Tso, G.K. *Predicting electricity energy consumption: a comparison of regression*

- analysis, decision tree and neural networks / G.K. Tso, K.K. Yau // *Energy*. – 2007. – Vol. 32, iss. 9. – P. 1761-1768.
75. Cao, J. Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks / J. Cao, X. Lin // *Energy Convers Manag.* – 2008. – Vol. 49, iss. 6. – P. 1396-1406.
76. Marquez, R. Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database / R. Marquez, C.F. Coimbra // *Sol Energy*. – 2011. – Vol. 85, iss. 5. – P. 746-756.
77. Short-term wind speed forecasting using support vector regression optimized by cuckoo optimization algorithm / J. Wang, Q. Zhou, H. Jiang, R. Hou // *Math Probl Eng.* – 2015. – Vol. 1. – P. 1-13.
78. Support-vector-machine-enhanced markov model for short-term wind power forecast / L. Yang, M. He, J. Zhang, V. Vittal // *IEEE Trans Sustain Energy*. – 2015. – Vol. 6, iss. 3. – P. 791-799.
79. Demand dispatch and probabilistic wind power forecasting in unit commitment and economic dispatch: a case study of Illinois / A. Botterud, Z. Zhou, J. Wang [et al.] // *IEEE Trans Sustain Energy*. – 2013. – Vol. 4, iss. 1. – P. 250-261.
80. Impact of public aggregate wind forecasts on electricity market outcomes / L. Exizidis, J. Kazempour, P. Pinson [et al.] // *IEEE Trans Sustain Energy*. – 2017. – Vol. 8, iss. 4. – P. 1394-1405.
81. Hyperparameter Search in Machine Learning [Electronic resource]. – Available from: <https://arxiv.org/abs/1502.02127>.
82. Bergstra, J. Random Search for Hyper-Parameter Optimization / J. Bergstra, Y. Bengio // *Journal of Machine Learning Research*. – 2012. – Vol. 13. – P. 281-305.
83. Overview of hyperparameter tuning [Electronic resource]. – Available from: <https://cloud.google.com/ai-platform/training/docs/hyperparameter-tuning-overview>.
84. Medium Term Load Forecasting for Jordan Electric Power System Using Particle Swarm Optimization Algorithm Based on Least Square Regression Methods / M. Hattab, M. Ma'itah, T. Sweidan [et al.] // *Journal of Power and Energy Engineering*.

- 2017. – Vol. 5, iss. 2. – P. 75-96.
85. Electricity load forecasting: a systematic review / I.K. Nti, M. Teimeh, O. Nyarko-Boateng, A.F. Adekoya // *Journal of Electrical Systems and Information Technology*. – 2020. – Vol. 7, iss. 1. – Art. 13.
 86. Abu-Shikhah, N. Medium-Term Electric Load Forecasting Using Multivariable Linear and Non-Linear Regression / N. Abu-Shikhah, F. Elkarmi, O. Aloquili // *Smart Grid and Renewable Energy*. – 2011. – Vol. 2, iss. 2. – P. 126-135.
 87. The impact of data filtration on the accuracy of multiple time-domain forecasting for photovoltaic power plants generation / S.A. Eroshenko, A.I. Khalyasmaa, D.A. Snegirev [et al.] // *Applied Sciences*. – 2020. – Vol. 10, iss. 22. – Art. 8265.
 88. Panapakidis, I.P. Implementation of modified versions of the K-means algorithm in power load curves profiling / I.P. Panapakidis, G.C. Christoforidis // *Sustainable Cities and Society*. – 2017. – Vol. 35. – P. 83-93.
 89. Clustering electricity usage profiles with K-means [Electronic resource]. – Available from: <https://towardsdatascience.com/clustering-electricity-profiles-with-k-means-42d6d0644d00>.
 90. Khalyasmaa, A.I. Inappropriate machine learning application in real power industry cases / A.I. Khalyasmaa, P.V. Matrenin, S.A. Eroshenko // *International Journal of Electrical and Computer Engineering*. – 2022. – Vol. 12, iss. 3. – P. 3023-3032.
 91. Matrenin, P.V. Improving accuracy and generalization performance of small-size recurrent neural networks applied to short-term load forecasting / P.V. Matrenin, V.Z. Manusov, A.I. Khalyasmaa // *Mathematics*. – 2020. – Vol. 8, iss. 12. – Art. 2169.
 92. Hartigan, J.A. Algorithm AS 136: A K-Means Clustering Algorithm / J.A. Hartigan, M.A. Wong // *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. – 1979. – Vol. 28, iss. 1. – P. 100-108.
 93. *Principal Manifolds for Data Visualization and Dimension Reduction* / A.N. Gorban, B. Kegl, D.C. Wunsch, A. Zinovyev. – Heidelberg: Springer, 2008. – 340 p.
 94. Adam: A method for stochastic optimization [Electronic resource]. – Available

from: <https://arxiv.org/pdf/1412.6980.pdf>

95. A Comprehensive Study on the Recent Progress and Trends in Development of Small Hydropower Projects / T.S. Kishore, E.R. Patro, V.S.K.V Harish, A.T. Haghighi // *Energies*. – 2021. – Vol. 14, iss. 10. – Art. 2882.
96. Bayazıt, Y. A study on transformation of multi-purpose dams into pumped storage hydroelectric power plants by using GIS model / Y. Bayazıt, R. Bakış, C. Koç // *International Journal of Green Energy*. – 2020. – Vol. 18, iss. 3. – P. 308-318.
97. Mayeda, A.M. Factors influencing public perceptions of hydropower projects: A systematic literature review / A.M Mayeda, A.D. Boyd // *Renewable and Sustainable Energy Reviews*. – 2020. – Vol. 121. – Art. 109713.
98. Jurasz, J. Solar–hydro hybrid power station as a way to smooth power output and increase water retention / J. Jurasz, B. Ciapała // *Solar Energy*. – 2018. – Vol. 173. – P. 675-690.
99. Efficiency evaluation of hydropower station operation: A case study of Longyangxia station in the Yellow River, China / J. Chang, Y. Li, M. Yuan, Y. Wang // *Energy*. – 2017. – Vol. 135. – P. 23-31.
100. Li, F-F. Multi-objective optimization for integrated hydro–photovoltaic power system / F-F. Li, Q. Jun // *Applied Energy*. – 2016. – Vol. 167. – P. 377-384.
101. Albo-Salih, H. Testing of an Optimization-Simulation Model for Real-Time Flood Operation of River-Reservoir Systems / H. Albo-Salih, L. Mays // *Water*. – 2021. – Vol. 13. – Art. 1207.
102. Vafakhahi, M. Application of artificial neural networks and adaptive neuro-fuzzy inference system models to short-term streamflow forecasting / M. Vafakhahi // *Canadian Journal of Civil Engineering*. – 2012. – Vol. 39, iss. 4. – P. 402-414.
103. Short-Term Power Generation Energy Forecasting Model for Small Hydropower Stations Using GA-SVM / G. Li, Y. Sun, Y. He [et al.] // *Mathematical Problems in Engineering*. – 2014. – Vol. 2014. – Art. 381387.
104. Applying a Correlation Analysis Method to Long-Term Forecasting of Power Production at Small Hydropower Plants / G. Li, C-X. Liu, S-L. Liao, C-T. Cheng // *Water*. – 2015. – Vol. 7, iss. 9. – P. 4806-4820.

105. Hinton, G.E. Connectionist learning procedures / G.E. Hinton // Artificial intelligence. – 1989. – Vol. 40. – P. 185-234.
106. Improving Regressors using Boosting Techniques [Electronic resource]. – Available from: https://www.academia.edu/28038936/Improving_regressors_using_boosting_techniques.
107. Breiman, L. Random Forests / L. Breiman // Machine Learning. – 2001. – Vol. 4. – P. 5-32.
108. Machine learning in Python [Electronic resource]. – Available from: <https://scikit-learn.org>.
109. Keras. Simple. Flexible. Powerful [Electronic resource]. – Available from: <https://keras.io/>.
110. Sachdev, H.S. Analysis and evaluation of small hydropower plants: A bibliographical survey / H.S. Sachdev, K.A. Ashok, N. Kumar // Renewable and Sustainable Energy Reviews. – 2015. – Vol. 51. – P. 1013-1022.
111. Algorithm for calculation and selection of micro hydropower plant taking into account hydrological parameters of small watercourses mountain rivers of Central Asia / M.S. Asanov, M.Kh. Safaraliev, T.Zh. Zhabudaev [et al.] // Int. J. Hydrogen Energy. – 2021. – Vol 46, iss. 75. – P. 37109-37119.
112. Recloser-Based Decentralized Control of the Grid with Distributed Generation in the Lahsh District of the Rasht Grid in Tajikistan, Central Asia / A. Ghulomzoda, A. Gulakhmadov, A. Fishov [et al.] // Energies. – 2020. – Vol. 13, iss. 14. – Art. 3673 (19 p.).
113. Антоненков, Д.В. Исследование ансамблевых и нейросетевых методов машинного обучения в задаче краткосрочного прогнозирования электропотребления горных предприятий / Д.В. Антоненков, П.В. Матренин // Электротехнические системы и комплексы. – 2021. – № 3(52). – С. 57-65.
114. XGBoost: A Scalable Tree Boosting System [Electronic resource]. – Available from: <https://arxiv.org/abs/1603.02754>.
115. Разработка моделей среднесрочного прогнозирования электропотребления в изолированно работающих энергосистемах на основе ансамблевых методов

- машинного обучения / С.М. Асанова, Дж.С. Ахъев, С.А. Дмитриев // Известия НТЦ единой энергетической системы. – 2021. – №1 (84). – С.32-39.
116. Séguin, S. Self-scheduling short-term unit commitment and loading problem / S. Séguin, P. Côté, Ch. Audet // IEEE Transactions on Power Systems. – 2016. – Vol. 31, iss. 1. – P. 133-142.
117. Generation Scheduling of a Hydrothermal System Considering Multiple Provincial Peak-Shaving Demands / J.J. Shen, Q.Q. Shen, S. Wang [et al.] // IEEE Access. – 2019. – Vol. 7. – P. 46225-46239.
118. Carpentier, P.L. Long-term management of a hydroelectric multireservoir system under uncertainty using the progressive hedging algorithm / P.L. Carpentier, M. Gendreau, F. Bastin // Water Resources Research. – 2013. – Vol. 49, iss. 5. – P. 2812-2827.
119. Optimal operation of interprovincial Hydropower System Including Xiluodu and Local Plants in Multiple Recipient Regions / J.J. Shen, X.F. Zhang, J. Wang [et al.] // Energies. – 2019. – Vol. 12, iss. 1. – Art. 144.
120. Self-scheduling of a hydro producer in a pool-based electricity market / A.J. Conejo, J.M. Arroyo, J. Contreras, F.A. Villamor // IEEE Trans. Power Syst. – 2002. – Vol. 17. – P. 1265–1272.
121. Tong, B. An MILP based formulation for short-term hydro generation scheduling with analysis of the linearization effects on solution feasibility / B. Tong, Q. Zhai, X. Guan // IEEE Trans. Power Syst. – 2013. – Vol. 28. – P. 3588–3599.
122. Zhao, T. Improved dynamic programming for hydropower reservoir operation / T. Zhao, J. Zhao, D. Yang // J. Water Resour. Plan. Manag. – 2012. Vol. 140. – P. 365–374.
123. Arce, A. Optimal dispatch of generating units of the Itaipú hydroelectric plant / A. Arce, T. Ohishi, S. Soares // IEEE Trans. Power Syst. – 2002. – Vol. 17. – P. 154-158.
124. Håberg, M. Fundamentals and recent developments in stochastic unit commitment / M. Håberg // Int. J. Electr. Power Energy Syst. – 2019. – Vol. 109. – P. 38-48.
125. Optimization of large-scale hydropower system peak operation with hybrid

- dynamic programming and domain knowledge / Z.-K. Feng, W.-J. Niu, C.-T. Cheng, X.-Y. Wu // *J. Clean. Prod.* – 2018. – Vol. 171. – P. 390-402.
126. Yu, B. Short-term hydro-thermal scheduling using particle swarm optimization method / B. Yu, X. Yuan, J. Wang // *Energy Convers. Manag.* – 2007. – Vol. 48. – P. 1902-1908.
127. A Lagrangian decomposition approach for the pump scheduling problem in water networks / B. Ghaddar, J. Naoum-Sawaya, A. Kishimoto [et al.] // *Eur. J. Oper. Res.* – 2015. – Vol. 241. – P. 490-501.
128. Petcharaks, N. Hybrid enhanced Lagrangian relaxation and quadratic programming for hydrothermal scheduling / N. Petcharaks, W. Ongsakul // *Electr. Power Compon. Syst.* – 2007. – Vol. 35. – P. 19-42.
129. Victorie, T. Unit commitment by a Tabu-search-based hybrid-optimization technique / T. Victorie, A. Jeyakumar // *IEE Proc. Gener. Transm. Distrib.* – 2005. – Vol. 152. – P. 563–574.
130. Nonlinear optimization method for short-term hydro scheduling considering head-dependency / J.P.S. Catalão, S.J.P.S. Mariano, V.M.F. Mendes, L.A.F.M. Ferreira, // *Eur. Trans. Electr. Power.* – 2008. – Vol. 20. – P. 172-183.
131. Adaptive ensemble models for medium-term forecasting of water inflow when planning electricity generation under climate change / P. Matrenin, M. Safaraliev, S. Dmitriev [et al.] // *Energy Reports.* – 2022. – Vol. 8. – P. 439-447.
132. Basics of evaluation methods of HPP resource usage during operation in IPS / A.V. Sidorova, A.Y. Arestova, A.G. Rusina, T.A. Filippova // *Advances in Engineering Research.* – 2017. – Vol. 133. – P. 810-815.
133. Manusov, V.Z. Robust Short-Term Wind Speed Forecasting Using Adaptive Shallow Neural Networks / V.Z. Manusov, E.A. Igumnova, P.V. Matrenin // *Problems of the Regional Energetics.* – 2020. – Vol. 47, iss. 3. – P. 68-79.
134. Ismail, M.S. Techno-economic analysis of an optimized photovoltaic and diesel generator hybrid power system for remote houses in a tropical climate / M.S. Ismail, M. Moghavvemi, T. M. I. Mahlia // *Energy conversion and management.* – 2013. – Vol. 69, – P. 163-173.

135. Ayodele, T.R. Increasing household solar energy penetration through load partitioning based on quality of life: The case study of Nigeria / T.R. Ayodele, A.S.O. Ogunjuyigbe // *Sustainable Cities and Society*. – 2015. – Vol. 18. – P. 21-31.
136. Ogunjuyigbe, A.S.O. Optimal allocation and sizing of PV/Wind/Split-diesel/Battery hybrid energy system for minimizing life cycle cost, carbon emission and dump energy of remote residential building / A.S.O. Ogunjuyigbe, T.R. Ayodele, O.A. Akinola. // *Applied Energy*. – 2016. – Vol. 171. – P. 153-171.
137. Furuholt, E. Life cycle assessment of gasoline and diesel / E. Furuholt // *Resources, Conservation and Recycling*. – 1995. – Vol. 14. – P. 251-263.
138. Fuel-cycle greenhouse gas emissions from alternative fuels in Australian heavy vehicles / T. Beer, T. Grant, D. Williams, W. Harry // *Atmospheric Environment*. – 2002. – Vol. 36. – P. 753–63.
139. Facanha, C. Evaluation of life-cycle air emission factors of freight transportation / C. Facanha, A. Horvath // *Environmental Science and Technology*. – 2007. – Vol. 41, iss. 20. – P. 7138-7144.
140. Integrated evaluation of greenhouse gas emissions (CO₂, CH₄, N₂O) from two farming systems in southern Germany / H. Flessa, R. Ruser, P. Dörsch [et al.] // *Agriculture Ecosystems & Environment*. – 2002. – Vol. 91. – P. 175-89.

СПИСОК РИСУНКОВ

- 1.1 Географическое расположение ИЭС ГБАО
- 1.2 Потенциал альтернативных источников энергии ГБАО
- 1.3 Графики генерации и потребления за 2020 г.
- 1.4 Однолинейная схема рассматриваемой изолированной энергосистемы ГБАО
- 3.1 Графики электропотребления за январь-март 2019 г.
- 3.2 Графики электропотребления за ноябрь-декабрь 2019 г.
- 3.3 Визуализация 4-х кластеров (а – данные за три дня, б – за четыре, в – за пять, г – за семь), цвет означает кластер
- 3.4 Визуализация 3-х кластеров при использовании данных за семь дней
- 3.5 Сравнение истинного и прогнозного графиков электропотребления
- 3.6 Графики притока воды 2010-2013 гг.
- 3.7 Графики притока воды 2015-2019 гг.
- 3.8 Матрица коэффициентов корреляции
- 3.9 Архитектура LSTM модели
- 3.10 Сопоставление результатов прогнозирования
- 3.11 Результаты для 2015 года (слева) 2019 (справа), зеленая линия – истинные значения, синяя – прогнозы АВ без самоадаптации (mRMSE 46.02 %, и 13.88 %), красная – с самоадаптацией (29.43 % и 12.62 %)
- 3.12 Фрагмент выборки данных (январь 2015–2019 гг.)
- 3.13 Матрица коэффициентов корреляции
- 3.14 Сравнение результатов моделей прогнозирования
- 3.15 Распределение дней по генерации и температуре, 2015 год слева, 2016 справа
- 3.16 Распределение дней по генерации и температуре, 2017 год слева, 2018 справа
- 3.17 Распределение дней по генерации и температуре, 2019 год слева, 2015-2019 гг. справа
- 3.18 Сопоставление прогнозов и истинных значений мощности, 2019 г. (тестовая часть выборки)
- 4.1 Графики электропотребления и генерации ноябрь 2015 – март 2016

- 4.2 Графики электропотребления и генерации ноябрь 2016 – март 2017
- 4.3 Графики электропотребления и генерации ноябрь 2017 – март 2018
- 4.4 Графики электропотребления и генерации ноябрь 2018 – март 2019
- 4.5 Графики электропотребления и генерации ноябрь 2019 – март 2020
- 4.6 Коэффициенты корреляции Спирмена
- 4.7 Точность прогноза разности потребления и выработки ГЭС и избыточное потребление ДЭС в тестовый период (ноябрь 2019 – март 2020)
- 4.8 Результаты прогноза ANN, январь – март 2020 г.
- 4.9 Результаты прогноза ANN, ноябрь – декабрь 2020 г.
- 4.10 Результаты прогноза AdaBoost, январь – март 2020 г.
- 4.11 Результаты прогноза AdaBoost, ноябрь – декабрь 2020 г.
- 4.12 Сопоставление прогнозов ANN и AdaBoost и истинным значением разности потребления и генерации ГЭС
- 4.13 Суммарные объемы генерации ДЭС в соответствии с планами, полученными от ANN и AdaBoost (шаг времени для суммирования – две недели).
- 4.14 Недельный график электропотребления для недели зимнего сезона
- 4.15 Недельный график электропотребления для недели весенне-осеннего периода
- 4.16 Графики выбросов ДЭС ГБАО для трех рассматриваемых сценариев

СПИСОК ТАБЛИЦ

- 1.1 Сведения о функционирующих ГЭС на территории ГБАО, находящихся на балансе «Pamir Energy»
- 1.2 Гидроэнергетические ресурсы ГБАО по категориям
- 1.3 Потенциальные запасы гидроэнергоресурсов ГБАО
- 1.4 Технические и экономические гидроэнергетические ресурсы малых водотоков ГБАО
- 1.5 Число дней без солнца в ГБАО
- 1.6 Суммарная месячная солнечная радиация в административном центре ГБАО Вт/м²
- 1.7 Среднегодовая скорость ветра по ГБАО
- 1.8 Установленные и располагаемые мощности ДЭС по Горно-Бадахшанским электрическим сетям (ЭС) на 26. 12. 1991 г.
- 2.1 Достоинства и недостатки методов среднесрочное прогнозирования электропотребления и генерации
- 3.1 Коэффициент корреляции между признаками и электропотреблением
- 3.2 Результаты прогнозных моделей с использованием метеорологических данных
- 3.3 Структура выборки, используемой при построении моделей
- 3.4 Используемые модели
- 3.5 Результаты LSTM, средняя ошибка в период 2015–2019 гг. 27.4 % или 4.93 м3/с
- 3.6 Результаты MLP, средняя ошибка в период 2015–2019 гг. 30.35 % или 5.45 м3/с без обучения онлайн и 25.32 % или 4.56 м3/с при онлайн обучении
- 3.7 Результаты AdaBoost [Decision Trees], средняя ошибка в период 2015–2019 гг. 24.43 % или 4.40 м3/с без обучения онлайн и 22.08 % или 3.97 м3/с при онлайн обучении

- 3.8 Результаты Random Forest, средняя ошибка в период 2015–2019 гг. 25.29 % или 4.55 м3/с без обучения онлайн и 22.63 % или 4.07 м3/с при онлайн обучении
- 3.9 Сопоставление результатов
- 3.10 Структура выборки для применения машинного обучения и пример заполнения
- 3.11 Размеры выборок при разных интервалах используемых ретроспективных данных
- 3.12 Результаты подбора гипер-параметров примененных моделей машинного обучения
- 3.13 Сравнение результатов моделей, приведены значения метрики ошибки MAPE, в каждой строке выделены наилучшие результаты, полученные для соответствующей модели
- 4.1 Структура датасета, отмеченные * признаки исключены после анализа
- 4.2 Результаты прогнозирования
- 4.3 Наилучшие результаты прогнозирования разности E и H .
- 4.4 Влияние ошибок прогноза на ДЭС
- 4.5 Наилучшие и наихудшие результаты каждой из моделей
- 4.6 Характерное число часов использования максимума нагрузки по дням недели
- 4.7 Результаты оценки степени экологичности и экономичности найденных планов генерации

ПРИЛОЖЕНИЕ

АКТ ВНЕДРЕНИЯ РЕЗУЛЬТАТОВ ДИССЕРТАЦИОННОЙ РАБОТЫ



Акт

О внедрении результатов кандидатской диссертационной работы

М.Х Сафаралиева

Настоящим актом подтверждается заинтересованность в использовании результатов и выводов диссертационной работы М.Х Сафаралиева на тему «Разработка моделей прогнозирования электропотребления и генерации ГЭС на среднесрочную перспективу в изолированных энергосистемах» по специальности 2.4.3. Электроэнергетика.

Результаты диссертационного исследования М.Х Сафаралиева могут быть использованы для обеспечения поддержания баланса генерации и потребления электроэнергии в изолированной энергосистеме ГБАО. Разработанный комплекс авторских моделей среднесрочного прогнозирования притока воды в водохранилища и генерации ГЭС обеспечивает оптимальный выбор состава генерирующего оборудования, определение перспективного запаса воды и оптимальное планирование графика сработки водохранилищ плотинных ГЭС. Следует отметить значимость разработанной модели комплексного планирования генерации, способной повысить эффективность использования ДЭС на ископаемом топливе, и, как следствие, снизить выбросы диоксида углерода в атмосферу и холостые сбросы запаса воды на плотинных ГЭС.

Технический директор

Шонизоров Ш.



ЦСК "Памирская Энергетическая Компания", Чумхурии Тоҷикистон, шаҳри Хоруғ, кӯчаи Гулмамадов 75
Тел.: +992 (3522) 2 28 10, 2 33 83, факс: + 992 (3522) 2 25 19
e-mail: info@pamirenergy.com / www.pamirenergy.com