УДК: 621.384.2 + 519.678.2 DOI: 10.26907/rwp29.2025.523-527

ИССЛЕДОВАНИЕ ТОЧНОСТИ ПРОГНОЗИРОВАНИЯ МПЧ КВ-КАНАЛОВ СВЯЗИ В ЗАВИСИМОСТИ ОТ СЕЗОННЫХ ИЗМЕНЕНИЙ И СДВИГОВЫХ ПРЕДИКТОРОВ

Н.А. Конкин, А.С. Елчанинова
Поволжский государственный технологический университет, 420000,
г. Йошкар-Ола, пл. Ленина, 3.
E-mail: konkina@volgatech.net

Аннотация. В работе исследовано влияние сезонных факторов, лаговых предикторов и форматов представления данных (h- и v-структуры) на точность прогнозирования максимальной применимой частоты (МПЧ) для КВ-связи с помощью рекуррентных нейросетей. Показано, что включение сдвиговых признаков (lag 1–3) и горизонтальной структуры данных обеспечивает более высокую точность в зимний и осенний периоды при сильной автокорреляции и низком уровне энтропии. Полученные результаты повышают надёжность оперативного планирования КВ-сеансов.

Ключевые слова: КВ-связь; Максимальная применимая частота (МПЧ); Рекуррентные нейросети (RNN, LSTM); Сезонные факторы; Лаговые предикторы; Форматы представления данных (h, v)

INVESTIGATION OF THE FORECAST ACCURACY OF HF COMMUNICATION CHANNEL MUF DEPENDING ON SEASONAL VARIATIONS AND LAGGED PREDICTORS

N.A. Konkin, A.S. Elchaninova

Abstract. The study examines the impact of seasonal factors, lagged predictors, and data representation formats (h- and v-structures) on the accuracy of forecasting the maximum usable frequency (MUF) for HF communication using recurrent neural networks. It demonstrates that incorporating lagged features (lags 1–3) and a horizontal data structure yields higher accuracy during winter and autumn when autocorrelation is strong and entropy is low, thereby enhancing the reliability of operational HF session planning.

Keywords: HF communication; Maximum Usable Frequency (MUF); Recurrent Neural Networks (RNN, LSTM); Seasonal factors; Lagged predictors; Data representation formats (h, v)

Введение

Ключевым параметром выбора диапазона рабочих частот ионосферной КВ-связи является максимальная применимая частота (МПЧ). Оперативное прогнозирование МПЧ критично для планирования КВ-сеансов, где допустима относительная погрешность прогноза не более 10%. Ионосферные параметры подвержены суточным и сезонным изменениям. Учёт сезонных факторов (например, индикаторов времени года) позволяет захватить годовые циклы, а включение лаговых (сдвиговых) предикторов – значений МПЧ за предыдущие интервалы времени – помогает моделировать автокорреляционные зависимости во временных рядах [1, 2].

В текущем исследовании при подготовке данных для прогнозных моделей используются разные форматы представления информации. Формат h (horizontal) описывает данные в виде горизонтальных временных срезов, фиксируя значения всех предикторов за один временной момент; он хорошо подходит для захвата моментальных характеристик системы. Формат v (vertical) организует данные как вертикальные временные срезы, где для каждого предиктора приводятся значения за последовательные временные точки. Этот подход позволяет модели учитывать динамику изменений и последовательные зависимости. Сравнение этих форматов важно для понимания того, какой из них лучше поддерживает обучение нейросетей с учётом сезонных и временных закономерностей.

Существующие методы прогноза МПЧ включают физические модели (например, ITU-R P.533) и статистические подходы [3]. Метод REC533 прогнозирует МПЧ через краткосрочное предсказание параметров foF2 и M3000F2 [4]. Среди статистических алгоритмов применяются автокорреляционные модели (точные на горизонтах до нескольких часов) и традиционные нейросети, использующие многодневные ряды ионосферных наблюдений. В современных исследованиях применяются также рекуррентные нейросети (RNN, в частности LSTM),

способные учитывать длительные последовательные зависимости. Однако физические модели не отражают локальных вариаций ионосферы, а статистические и нейросетевые методы требуют больших объёмов данных, чтобы адаптироваться к резким сезонным изменениям.

Цель настоящего исследования – проанализировать влияние сезонных факторов, сдвиговых предикторов и форматов данных (h и v) на точность моделей прогноза МПЧ, обученных с помощью рекуррентных нейросетей. Научная новизна заключается в системном учёте сезонности, временных лагов и форматов представления данных при обучении RNN-моделей. Практическая значимость работы состоит в повышении надёжности прогнозов ионосферных условий и оптимизации планирования КВ-сеансов связи.

Алгоритм исследования

На рисунке 1 представлены основные алгоритмы исследования.

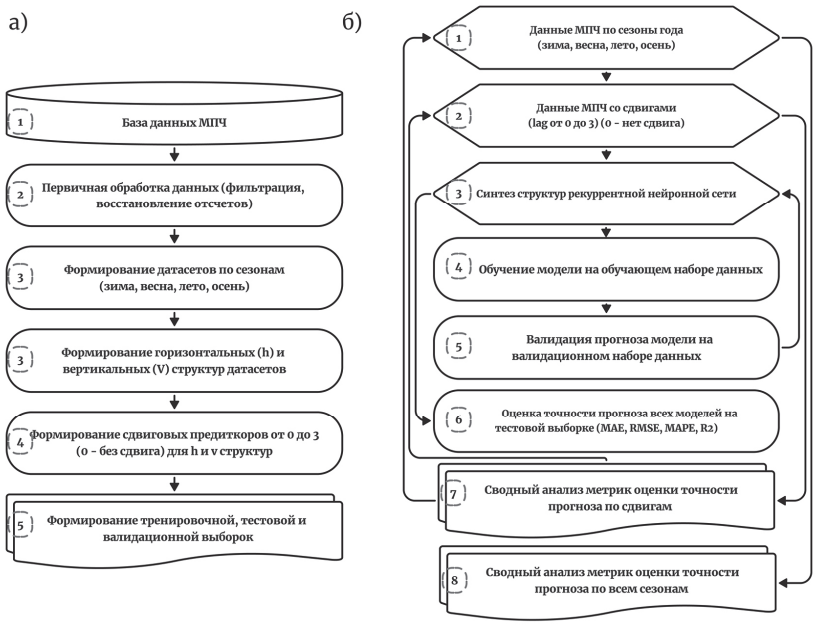

Рис. 1. Алгоритмы исследований: а) блок-схема формирования датасетов в зависимости от сезона года, формата данных и количества сдвигов, б) блок-схема проведенного эксперимента

Подготовка данных для прогнозирования максимальной применимой частоты (МПЧ) осуществляется поэтапно (рис. 1а). На шаге 1 формируется исходная база данных, содержащая измеренные значения МПЧ КВ-каналов за продолжительный временной интервал. Эти данные служат основой для анализа временных закономерностей и сезонных изменений. На шаге 2 проводится первичная обработка данных, включающая фильтрацию шумов и выбросов, а также восстановление недостающих временных отсчётов с целью формирования непрерывного и целостного временного ряда. Это критически важно для обеспечения корректного обучения моделей, чувствительных к пропускам в данных. На шаге 3 выполняется формирование структур данных в двух форматах: горизонтальном (h), где фиксируются значения всех предикторов на одном временном шаге, вертикальном (v), где для каждого предиктора строится последовательность значений за несколько предыдущих временных шагов (сдвиговые предикторы). Эти форматы позволяют модели учитывать как статические характеристики момента, так и динамические зависимости во времени. На шаге 4 формируются датасеты с различными сдвигами (лагами) – от 0 (без лагов) до 3 предыдущих временных шагов, что позволяет исследовать влияние временной памяти на точность прогноза. Временной шаг (lag) составляет 1 день с шагом 5 минут (288 отсчетов). На шаге 5 данные делятся на тренировочную, валидационную и тестовую выборки. Размер тренировочной выборки 18 дней с шагом 5 минут, валидационной – 1 день, тестовой – от 1 до 7 дней.

На рисунке 1б представлен алгоритм проведенного исследования на основе данных алгоритма 1а. Этапы 1 и 2 алгоритма исследования точности прогнозирования МПЧ КВ-каналов формируют конкретный датасет для машинного обучения по каждому сезону года со

сдвигами от 0 до 3. На основе подготовленных данных разрабатываются архитектуры рекуррентных и полносвязных нейронных сетей (LSTM-MLP). Выполняется. настройка числа скрытых слоёв, нейронов, функций активации и других гиперпараметров (этап 3). Модели обучаются на тренировочных наборах данных, с использованием разработанных архитектур RNN (этап 4). Валидация выполняется на выделенной валидационной выборке, что позволяет контролировать качество обучения и избегать переобучения (этап 5). Окончательная оценка точности прогнозов проводится на тестовой выборке. Рассчитываются стандартные метрики: MAE (средняя абсолютная ошибка), RMSE (среднеквадратичная ошибка), MAPE (средняя абсолютная процентная ошибка), R^2 (коэффициент детерминации) (этап 6). На этапах 7 и 8 выполняются сводный анализ по сдвигам и сезонам соответственно. Исследование по данному алгоритму (рис. 1б) выполняется отдельно для форматов v и h.

Результаты исследования

По результатам эксперимента получены усредненные оценки ошибок в при прогнозировании МПЧ по относительной метрике MAPE (Mean Absolute Percentage Error (средняя абсолютная процентная ошибка) и абсолютной метрике RMSE (Root Mean Square Error (среднеквадратичная ошибка) за 7 дней для различных сдвиговых признаков, охватывающие зимний (февраль), весенний (май), летний (август) и осенний (ноябрь) периоды.

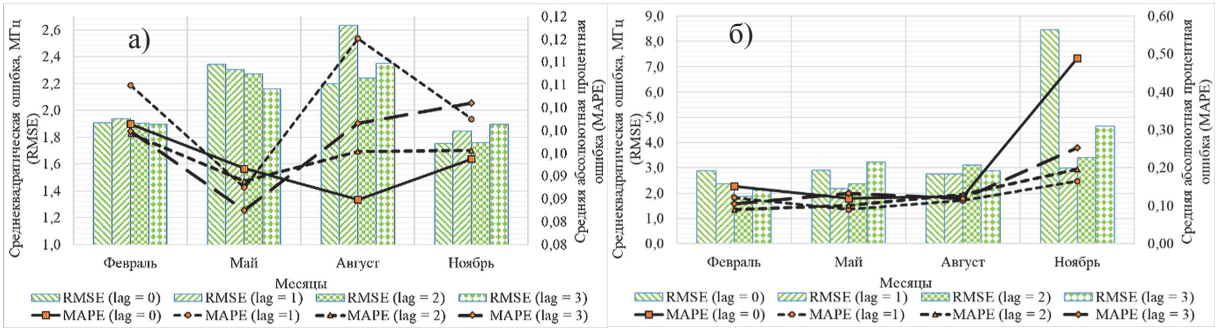

Рис. 2. Ошибки прогнозирования МПЧ по метрикам RMSE и MAPE для h- (а) и v-структуры (б) данных с различными сдвиговыми признаками

Метрики прогноза (рис. 2) показывают, что для h-структуры наименьшая усредненная ошибка прогнозирования МПЧ в феврале по RMSE составляет 1,896 МГц и MAPE 0,100 при сдвиге на 3 дня, в мае по RMSE 2,161 МГц и MAPE 0,086 при сдвиге на 3 дня, в августе по RMSE составляет 2,239 МГц и MAPE 0,095 при сдвиге на 3 дня, в ноябре по RMSE составляет 1,760 МГц и MAPE 0,097 при сдвиге на 2 дня.

Для v-структуры наименьшая усредненная ошибка прогнозирования МПЧ в феврале по RMSE составляет 1,868 МГц и MAPE 0,090 при сдвиге на 3 дня, в мае по RMSE 2,183 МГц и MAPE 0,090 при сдвиге на 1 день, в августе по RMSE составляет 2,761 МГц и MAPE 0,114 при сдвиге на 1 день, в Ноябре по RMSE составляет 2,995 МГц и MAPE 0,165 при сдвиге на 2 дня.

Результаты анализа указывают на то, что применение сдвиговых признаков приводит к повышению точности прогнозирования МПЧ. Это обусловлено тем, что сдвиговые признаки (lag) помогают модели учитывать автокорреляционный характер данных, т.е. зависимость текущих данных от предыдущих. Т.к. при использовании сдвига на 1, 2, 3 дня модель способна учитывать информацию о предыдущих днях и наблюдается повышение точности. Однако в случае, если модель начинает учитывать информацию, которая не имеет значения для текущего прогноза или существует слабая зависимость между текущими и предыдущими значениями, может наблюдаться ухудшение точности.

С целью выявления зависимости точности прогнозирования от сезонности введем параметр PACF (Partial Autocorrelation Coefficient) и ApEn (Approximate Entropy). Параметр ApEn позволяет проанализировать предсказуемость и сложность временного ряда. Низкие значения ApEn указывают на высокую упорядоченность данных и предсказуемость, тогда как

высокие значения свидетельствуют о большей хаотичности и сложности данных. Также значения ApEn меньше 0,1 указывают на стационарность данных, при диапазоне от 0,1-02 - средняя сложность данных из-за наличия нелинейных эффектов, при ApEn больше 0,2 - данные относят к сложным и нестационарным.

Что касается параметра PACF, он помогает измерять степень корреляции между наблюдениями временного ряда на различных лагах. Для оценки корреляции выделяют 4 диапазона: при PACF меньше 0,1 корреляция отсутствует, при диапазоне от 0,1–0,3 присутствует слабая корреляция; в диапазоне 0,3–0,5 наблюдается средняя корреляция; при значениях больше 0,5 наблюдается сильная корреляция. Сводные результаты по эксперименту представлены в таблице 1.

Таблица 1. Сводная таблица по метрикам датасета PACF, ApEn и прогноза RMSE в зависимости от сезона и типа структуры датасета

Сезон	Тип структуры	Вариация PACF	ApEn	Усредненное RMSE (МГц)
Зима	h-структура	0,604–0,804	0,095–0,096	1,896–1,938
	v-структура	0,532–0,791		1,868–2,761
Весна	h-структура	0,009–0,029	0,096–0,097	2,161–2,343
	v-структура	0,009–0,030		2,183–3,244
Лето	h-структура	0,009–0,053	0,162–0,166	2,191–2,634
	v-структура	0,005–0,051		2,761–3,113
Осень	h-структура	0,581–0,956	0,952	1,750–1,896
	v-структура	0,185–0,830		2,995–8,460

Таким образом, анализируя параметры PACF и ApEn и усредненную метрику RMSE было выявлено, что для h-структуры в зимний и осенний период модель обладает более высокой точностью по сравнению с летним и весенним периодом. Поскольку модель обладает более высокой автокорреляцией и нестационарностью данных, наблюдаются более низкие значения по усредненной метрике RMSE. Для v-структуры более высокую точность модель продемонстрировала в зимний период. В весенний и летний из-за слабой корреляции и более сложного временного ряда данных, модель обладает более высоким значением ошибки. В осенний период модель при lag=1 показала низкую точность из-за отсутствия корреляции между данными. Также было установлено, что параметр PACF и точность сдвиговых предикторов (lag) взаимосвязаны, то есть, чем выше значения частично автокорреляционной функции, тем выше точность модели при применении сдвигов. И наоборот, чем ниже значения PACF, тем ниже точность модели при использовании сдвиговых признаков.

Заключение

В настоящей работе была проведена всесторонняя оценка влияния сезонных факторов, сдвиговых предикторов и форматов представления данных (h- и v-структур) на точность прогноза максимальной применимой частоты (МПЧ) КВ-связи с использованием рекуррентных нейросетей.

В ходе эксперимента определено, что для оперативного прогнозирования МПЧ целесообразно применять RNN-модели, обученные на горизонтальной структуре данных с включением сдвиговых признаков (предпочтительно lag = 2–3), особенно в периоды высокой автокорреляции ионосферных параметров (зима и осень). Полученные результаты свидетельствуют о практической значимости системного учёта сезонности и лаговых зависимостей для повышения надёжности КВ-сеансов.

Благодарности

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 23-19-00145).

Список литературы

1. Прогнозирование полосы прозрачности ионосферной линии КВ-радиосвязи. – Ч. 1. Теоретико-экспериментальные основы нового когнитивного подхода к диагностике и прогнозированию максимально применимой частоты / Д.В. Иванов, В.А. Иванов, Н.В. Рябова [и др.] // Радиотехника. – 2025. – Т. 89. – № 2. – С. 152–166. – https://doi.org/10.18127/j00338486-202502-20. – EDN ENIEKQ.
2. Подходы к экспериментальному оцениванию SDR сенсором вариаций спектральной плотности мощности помех и прогнозированию вариаций их среднего уровня при изменении полосы КВ-радиоканала / Д.В. Иванов, В.А. Иванов, Н.В. Рябова, Р.Р. Бельгибаев // Электромагнитные волны и электронные системы. – 2025. – Т. 30. – № 1. – С. 68–78. – https://doi.org/10.18127/j5604128-202501-07. – EDN IDTKHV.
3. Международный союз электросвязи (МСЭ). Рекомендация ITU-R P.533-12: Метод прогнозирования характеристик коротковолновых (КВ) радиолиний. – Женева: МСЭ, 2013. – URL: https://www.itu.int/rec/R-REC-P.533-12-201309-S
4. Международный союз электросвязи. Рекомендация ITU-R P.533-12: Метод прогнозирования характеристик коротковолновых (КВ) радиолиний. – Женева: МСЭ, 2013. – 66 с. – URL: https://www.itu.int/rec/R-REC-P.533-12-201309-S