

ИНТЕГРАЦИОННОЕ ПРОГНОЗИРОВАНИЕ НЕСТАЦИОНАРНЫХ ПРОЦЕССОВ, ПРЕДСТАВЛЕННЫХ ВРЕМЕННЫМИ РЯДАМИ. ОБЗОР¹

Авдеева З. К.², Коврига С. В.³
(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Обзор охватывает основные направления и подходы к интеграционному прогнозированию нестационарных процессов, представленных временными рядами. Ключевым источником порождения нестационарности являются быстрые и плохо предсказуемые изменения внешней среды, под влиянием которых происходят структурные сдвиги в процессах, протекающих в сложных экономических и социально-политических системах. Решение задач прогнозирования динамики таких объектов в контексте повышения точности прогноза усложняется по мере увеличения горизонта прогнозирования, что обуславливает потребность в моделях и методах, способных обрабатывать разнородную информацию. Интеграционные методы – это методы, позволяющие наряду с количественными данными учитывать суждения (прогнозистов, экспертов, аналитиков) и информацию из разнородных информационных источников на разных этапах решения задачи и тем самым прямым или косвенным способом включать их в формируемый прогноз. Развитие таких методов направлено на повышение точности прогноза через использование всей доступной информации об объекте прогнозирования, включая данные об эндогенных и экзогенных факторах влияния на него. В обзоре внимание было сконцентрировано на современном состоянии в области интеграционного прогнозирования, на существующих проблемах и путях их решения.

Ключевые слова: нестационарные процессы, временные ряды, прогнозирование, суждения.

1. Введение

Анализ и прогнозирование временных рядов является популярным прогностическим инструментом для работы с большими объемами данных, отражающих закономерности поведе-

¹ Работа выполнена при частичной финансовой поддержке Российского научного фонда (грант № 23-21-00455).

² Зинаида Константиновна Авдеева, к.т.н., с.н.с. (avdeeva@ipu.ru).

³ Светлана Вадимовна Коврига, н.с. (kovriga@ipu.ru).

ния исследуемых процессов в экономических (планирование на разных уровнях управления, товарные и финансовые рынки, управление цепочками поставок и др.), социально-политических (протестная активность, избирательные компании и др.), соcio-технических системах (социальные сети) [4, 6].

Несмотря на то, что количественные методы анализа и прогнозирования временных рядов постоянно развиваются, их потенциала недостаточно для эффективного прогнозирования нестационарных процессов, описывающих динамику таких систем. Одним из ключевых источников порождения нестационарности являются быстрые, плохо предсказуемые изменения внешней среды, под влиянием которых происходят структурные сдвиги и структурные изменения прогнозируемых процессов.

При этом проблема точности прогноза нарастает по мере увеличения горизонта прогнозирования. Прогнозирование временного ряда на несколько шагов вперед на длительный период времени является сложной задачей, поскольку уровень неопределенности прогноза увеличивается с ростом горизонта [34].

Многочисленные исследования показывают, что количественные (статистические) и качественные (основанные на экспертных знаниях) методы имеют ограничения и проблемы с точностью. Оба класса методов зависят от качества исходных данных и знаний, от типа неопределенности внешних факторов, от волатильности значимых параметров и пр., и, следовательно, от способности улавливать в информационном пространстве сигналы, указывающие на начало значимых изменений (см., например, [39, 46]).

Такие сигналы невозможно идентифицировать исключительно на основе анализа количественных данных, поскольку они отсутствуют в данных временных рядов. Необходимо привлечение экспертных знаний и выбор подходящих методов обработки этих знаний в контексте решения задачи прогнозирования. Поэтому в условиях высокой изменчивости факторов воздействия и самого прогнозируемого процесса, когда статистические методы теряют свою эффективность, повышение качества прогноза по мере роста горизонта прогнозирования воз-

можно путем сочетания качественных и количественных методов.

Основная цель данной статьи – представить обзор основных направлений и подходов к прогнозированию нестационарных процессов с прямым и косвенным включением в формируемые прогнозы суждений (прогнозистов, экспертов и аналитиков¹) и информации из разнородных информационных источников. Рассматривается прогнозируемый объект, процесс изменения которого описывается временным рядом или группой временных рядов, динамика которых может изменяться не только вследствие эндогенных изменений, но и под воздействием экзогенных изменений в окружающей среде объекта. Под влиянием экзогенных факторов в процессе могут возникать структурные сдвиги: скачки уровня или волатильности, изменения трендов временных рядов, изменения связей между компонентами рядов. Факторы влияния могут быть описаны 1) временными рядами; 2) данными, полученными в результате структурирования и обработки различных источников информации о состоянии и направлении влияния факторов на текущее состояние объекта прогнозирования.

Проведенный в рамках обзора анализ подходов к прогнозированию такого рода объектов позволяет выделить общую тенденцию развития методов прогнозирования временных рядов. За период начиная с основополагающей работы [20] и вплоть до наших дней широкий круг специалистов подтверждает, что статистические комбинации прогнозов повышают точность прогнозов. Наряду с этим эмпирические данные свидетельствуют об усилении тенденций, связанных с развитием комбинированных методов по пути 1) сочетания статистических прогнозов и прогнозов, полученных с помощью машинного обучения [45], и 2) включения суждения в методы формирования прогнозов [43, 45]. При этом также отмечается необходимость повышения

¹ Прогнозист – специалист по прогнозированию, носитель знаний о моделях и методах прогнозирования; эксперт, аналитик – носитель знаний предметной области (допускается, что специалист может быть носителем и тех и других знаний).

интерпретируемости результатов прогнозов, чтобы они были пригодными для принятия решений.

С учетом ориентации на объекты прогнозирования – нестационарные процессы, порождаемые изменениями внешней среды, а также с учетом современных тенденций построения прогнозов на основе комбинирования методов, в том числе развития интеграционных методов, позволяющих включать суждения в формируемый прогноз, в обзоре внимание было сконцентрировано на современном состоянии в этой области, в большей степени отражающих возможности и ограничения интеграционного прогнозирования, использующего разнотипную информацию, а также на нерешенных проблемам. При этом в качестве основных критериев применимости моделей и методов к прогнозированию нестационарных процессов выделены точность формируемого прогноза, интерпретируемость его результатов для принятия решений, своевременность получения прогноза.

Насколько известно авторам, в отечественной публикационной среде обзор, систематизирующий основные направления и подходы к развитию интеграционных методов прогнозирования временных рядов, не представлен. Существующие обзоры, известные авторам, посвящены количественному прогнозированию временных рядов (например, один из последних обзоров представлен в [1]). Зарубежные обзоры, посвященные интеграционному прогнозированию, как правило, затрагивают отдельные аспекты проблемы включения суждений в формируемые прогнозы, в частности, рассматривают вопросы понимания и оценки основных когнитивных особенностей и поведенческих проблем прогнозистов, которые необходимо учитывать при разработке интеграционных методов, применимость таких методов в некоторой прикладной области (например, прогнозирования спроса на продукцию/продаж в контексте цепочки поставок и т.д.), сравнение эффективности оценочных и статистических методов (см. например, недавние обзоры [15, 32, 43]) и др.

Раздел 2 включает вступительную часть к обзору. В разделе 3 представлено общее описание особенностей и подходов к формированию прогнозов с использованием количественных

методов как необходимой базы для решения задач прогнозирования нестационарных процессов в условиях структурных сдвигов. Раздел 4 содержит обзор основных направлений и подходов к интеграционному прогнозированию, использующему суждения и количественные данные на разных этапах прогнозирования, выделены проблемы разработки и применения таких методов и пути их решения. В заключении представлены обобщающие выводы.

2. Общие положения

При выделении основных направлений и подходов к прогнозированию нестационарных процессов со структурными сдвигами ориентация была на следующие основные классификационные признаки: степень формализации, общий принцип действия, способ получения прогнозной информации, горизонт прогноза, тип информации и данных о процессе.

Следует отметить, что рассмотренные направления и подходы к прогнозированию процессов, представленных временными рядами, в большей степени являются общими, независимо от их стационарности или нестационарности. Специфика, как правило, связана с применением конкретных моделей и методов прогнозирования нестационарных процессов. Выбор метода прогнозирования зависит от ряда факторов: свойств прогнозируемого ряда (случайное блуждание или фрактальный процесс); регулярности решения задачи (одноразовый или повторяющийся); формы представления прогноза (конкретное значение, конкретное значение и/или интервал, направление тренда – восходящий, нисходящий – и оценка его продолжительности), длины горизонта прогноза; типа информации и данных о прогнозируемом процессе, их доступности и полноте; требований к точности прогноза [1].

По типу используемых данных и информации и степени формализации все методы прогнозирования делятся на 1) интуитивные (экспертные), 2) формализованные (количественные) и 3) интеграционные (смешанные), использующие 1) и 2).

На рис. 1 отражена применимость различных классов методов на горизонте прогнозирования в зависимости от обеспеченности информацией об объекте прогнозирования.

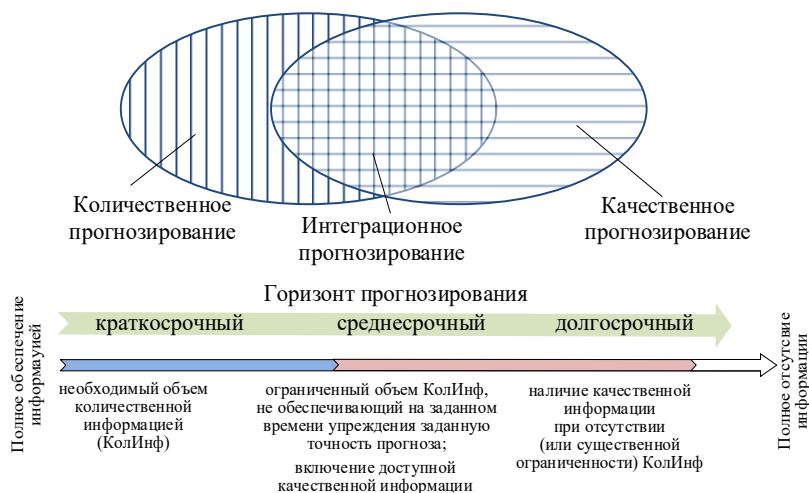


Рис. 1. Применимость методов прогнозирования в зависимости от обеспеченности информацией на горизонте прогнозирования

Класс интуитивных (экспертных) методов применяется в случаях, когда объект прогнозирования либо слишком прост, либо настолько сложен, что аналитически учесть влияние многих факторов практически невозможно [4], либо прогноз формируется на слишком большом горизонте прогнозирования (как правило, это долгосрочное прогнозирование).

В результате опроса экспертов полученные экспертные оценки используют как конечные прогнозы. В зависимости от общих принципов действия эти методы делятся на две группы: индивидуальные и коллективные экспертные оценки. К типовым экспертным методам, применяемых для прогнозирования развития таких объектов, относятся метод интервью, метод сценариев, метод Дельфи [45].

В обзоре подробно не представлен класс чисто экспертных методов, так как рассматривается объект прогнозирования, ди-

намика которого представлена временным рядом. В основе анализа и прогнозирования временных рядов, как правило, лежат методы формальной обработки. Включение в обзор общего описания особенностей и подходов к формированию прогнозов на основе количественных методов обусловлено тем, что они рассматриваются как необходимая методологическая база для решения задач прогнозирования нестационарных процессов со структурными сдвигами на горизонте прогнозирования.

Далее более подробно рассмотрим общие подходы к прогнозированию нестационарных процессов на основе количественных и интеграционных методов.

3. Прогнозирование на основе количественной информации

Количественные методы по общему принципу действия можно разделить на единичные (когда прогноз формируется на основе единственного метода) и комбинированные (на основе интеграции нескольких методов).

Комбинирование прогнозов зарекомендовало себя в качестве эффективного подхода к прогнозированию, приводящего в основном к повышению точности прогнозов [5, 8, 14, 45], в том числе в задачах обнаружения структурных сдвигов временных рядов, обусловленных изменяющимися во времени внешними условиями. При этом изменения динамики могут быть вызваны не только постепенным изменением воздействия внешних факторов, но и внезапными событиями [18]. Как отмечалось во введении, наблюдается усиление тенденции развития комбинированных методов.

Актуальность совместной применимости методов для комбинированных прогнозов (независимо от используемых методов: только количественных или интеграционных) следует из следующих условий [14]:

- если нет уверенности, какой метод прогнозирования является наиболее точным. Это может быть вызвано разными причинами: новой ситуацией, наличием разнородного набора временных рядов или ожиданием того, что будущее будет особенно

нестабильным, изменчивым. Даже если можно определить лучший метод, объединение может быть полезным, если другие методы дают некоторую дополнительную информацию;

- когда нет уверенности в ситуации прогнозирования. В ситуациях неопределенности объединение может уменьшить ошибку. Неопределенность ситуации может быть оценена на основании необъяснимых значительных отклонений в прошлом или ожидаемых неустойчивых изменений в будущем. Поскольку неопределенность увеличивается с увеличением горизонта прогноза, комбинирование должно быть особенно полезным для среднесрочных и долгосрочных прогнозов;

- когда важно избежать больших ошибок.

Базовыми принципами объединения прогнозов является использование [7, 8, 14]

- нескольких источников данных, позволяющих добавить дополнительную полезную информацию, способствующую более полному и точному описанию прогнозируемого процесса, а также скорректировать предвзятость (при привлечении экспертов);

- разных методов или данных, или сочетаний различных методов и данных;

- формальных процедур для объединения прогнозов;

- знаний предметной области, чтобы варьировать весами методов;

- равных весовых коэффициентов, если нет веских доказательств в поддержку неравного веса прогнозов;

- бóльшего веса для метода, который априори известен как лучший по сравнению с остальными.

Линейная форма является наиболее распространенной формой при объединении прогнозов в силу своей простоты и удобства в использовании [8]:

$$(1) \quad \hat{F}_{fn} = \sum_{i=1}^n w_i \hat{F}_i,$$

где $\hat{F}_{fn}$ – объединенный прогноз; $\hat{F}_i$ – i -й частный прогноз; w_i – весовой коэффициент, с которым i -й частный прогноз был включен в объединенный прогноз.

Одной из важнейших проблем объединения прогнозов является нахождение весовых коэффициентов, с которыми частные прогнозные значения объединяются в общем прогнозе. От выбора метода построения весовых коэффициентов зависит уровень повышения точности прогнозирования. На сами весовые коэффициенты накладываются только два ограничения: сумма всех весовых коэффициентов должна быть равно единице и весовые коэффициенты должны находиться в интервале $[0, 1]$ [8].

В [8] рассмотрены основные требования, предъявляемые к выбору весов, преимущества и недостатки различных подходов к их определению. В недавнем обзоре [56] можно ознакомиться с современным состоянием в области комбинирования прогнозов (независимо от того, объединяются только количественные прогнозы или, в том числе, с включением суждений), в части определения потенциала и ограничений различных стратегий комбинирования (выбора весов объединяемых прогнозов).

Также поднимается проблема выбора из пула прогнозов до комбинирования и отмечается необходимость разработки эмпирических рекомендаций и количественных метрик, которые помогут прогнозистам в выборе [55, 56].

По способу обработки и получения прогнозной информации можно выделить три основных подхода [43]: 1) экстраполяция, 2) многомерные методы, 3) компьютероёмкие методы.

Экстраполяция использует исторические данные для прогнозирования. Из экстраполяционных методов экспоненциальное сглаживание, которое усредняет (сглаживает) данные временных рядов, является одним из широко практикуемых классических методов. Авторегрессионное интегрированное скользящее среднее (ARIMA) и его многочисленные варианты (например, ARIMAX, ARMA-GARCH, ARFIMA) также являются популярными экстраполяционными методами, поскольку они могут учитывать тенденции, сезонность, ошибки и нестационарные аспекты временного ряда.

Многомерные методы являются формами регрессионного анализа и предполагают наличие причинной связи между зависимой переменной и одной или несколькими независимыми (объясняющими) переменными. Такие методы способны решить

проблему среднесрочного прогнозирования нестационарных временных рядов, рассматривая различные экзогенные переменные, которые могут повлиять на объект прогнозирования.

Компьютероемкие (computer intensive) методы¹ прогнозирования являются более поздними разработками. Недавний рост их популярности стимулирован ростом интеллектуального анализа данных и прогнозной аналитики, а также ростом накопленных количественных массивов нужной длины и качества. Еще одной причиной их развития обусловлено тем, что традиционные методы многомерного прогнозирования ограничены задачами низкой размерности, линейными зависимостями и короткими горизонтами. Технологические достижения вместо этого смещают акцент на проблемы, характеризующиеся большим количеством переменных, нелинейными зависимостями и длительными горизонтами прогнозирования. За последние несколько лет большинство наиболее эффективных методов многомерного прогнозирования были основаны на моделях глубокого обучения. Однако такие модели характеризуются высокими требованиями к доступности данных и вычислительным ресурсам и страдают от недостаточной интерпретируемости [25].

Также по способу обработки информации с учетом горизонта прогнозирования можно выделить два стандартных подхода: прогнозирование на один шаг и на несколько шагов вперед.

Последовательное применение одношаговой прогнозной модели, учитывающей взаимосвязи между последовательными значениями процесса, может быть полезен только на ограниченном интервале, поскольку не учитывает изменения, возникающие на горизонте прогнозирования. Если состояние объекта зависит не только от значений процесса, но и от значений упреждающих факторов, описываемых временными рядами других процессов, то использование информации о состоянии упреждающих факторов позволяет повысить качество прогноза.

Многошаговое прогнозирование ориентировано на формирование средне- и долгосрочных прогнозов.

¹ Это методы, требующие значительных вычислительных ресурсов и времени.

В общем виде задача многошагового прогнозирования

– для одномерного исторического временного ряда, состоящего из T наблюдений, состоит в прогнозировании следующих H значений, где $H > 1$ – горизонт прогноза:

$$(2) \{y_1, y_2, \dots, y_T\} \Rightarrow \{y_{T+1}, y_{T+2}, \dots, y_{T+H}\}.$$

– для многомерного ряда, состоящего из N рядов исторических наблюдений вида $\{y_1^k, y_2^k, \dots, y_T^k\}$, –

$$(3) \{Y_T^1, \dots, Y_T^k, \dots, Y_T^N\} \Rightarrow \{Y_{T+H}^1, \dots, Y_{T+H}^k, \dots, Y_{T+H}^N\},$$

где $Y_{T+H}^k = \{y_{T+1}^k, y_{T+2}^k, \dots, y_{T+H}^k\}$.

Отличие от прогнозирования на один шаг вперед, прогнозирования состояния объекта (например, цен на товарных рынках) на заданный горизонт H с пошаговым разбиением представляет трудную задачу; качество и методы решения зависят от свойств прогнозируемого объекта. Также сложность обусловлена тем, что приходится иметь дело с различными проблемами: накоплением ошибок, снижением точности и увеличением неопределенности [38, 52]. Однако задачи прогнозирования на несколько шагов все больше находят распространение в различных приложениях.

Существует три основные стратегии решения задачи многошагового прогнозирования: рекурсивная, прямая (независимая) и совместная стратегии [52, 53].

В рекурсивной стратегии (Recursive, S^{Rec}) формирование прогноза $\hat{y}_{T+h}$ на h -шагов вперед происходит путем итеративного вычисления значений в $t+1$, $t+2$, ... по сформированной одношаговой модели прогноза $M^{one-step}$:

$$(4) M^{one-step}: \hat{y}_{t+1} = f(y_t, y_{t-1}, \dots, y_{t-d+1}) + w,$$

где $t \in d, \dots, T-1$; d – количество прошлых наблюдений; w – слагаемое, включающее ошибку модели и внешние возмущения/шум;

$$(5) S^{Rec}: \hat{y}_{T+h} = \begin{cases} \hat{f}(y_T, \dots, y_{T-d+1}), & \text{если } h=1; \\ \hat{f}(\hat{y}_{T+h-1}, \dots, \hat{y}_{T+1}, y_T, \dots, y_{T-d+1}), & \text{если } h \in \{2, \dots, d\}; \\ \hat{f}(\hat{y}_{T+h-1}, \dots, \hat{y}_{T+h-d}), & \text{если } h \in \{d+1, \dots, H\}. \end{cases}$$

В зависимости от шума, присутствующего во временных рядах и горизонте прогнозирования, рекурсивная стратегия может быть недостаточно эффективной в задачах многошагового прогнозирования. Эта стратегия чувствительна к накоплению ошибок на горизонте прогнозирования.

Прямая стратегия ($\text{Direct}, S^{\text{Dir}}$) заключается в формировании многомасштабного прогноза $\hat{y}_{T+h}$ на горизонте H , когда исходный временной ряд преобразуется в ряды в h -временных масштабах, например, для годового прогноза в четырёх масштабах: месяц, квартал, полгода, год, а далее формируется h прогнозных моделей:

$$(6) \quad M^h: y_{t+h} = f_h(y_t, \dots, y_{t-d+1}) + w,$$

где $t \in \{d, \dots, T-H\}$, $h \in \{1, \dots, H\}$.

В этой стратегии прогноз есть h условно независимых прогнозов

$$(7) \quad S^{\text{Dir}}: \hat{y}_{T+h} = \hat{f}_h(y_T, \dots, y_{T-d+1}).$$

S^{Dir} не использует приближительных значений для расчета прогнозов, что делает ее невосприимчивой к накоплению ошибок. Такая стратегия применяется и для улучшения качества краткосрочных прогнозов, и для формирования среднесрочных прогнозов. При этом эта стратегия требует большого вычислительного времени, поскольку необходимо построить h моделей. Сочетание многомасштабного прогноза с комбинированием результатов прогноза при поступлении новых данных обеспечивает возможность коррекции за счет выбора лучших из имеющихся наборов. Однако предположение о независимости h прогнозов между собой не всегда верно, так как есть методы, которые показывают сложные зависимости между поведением ряда в разных масштабах.

Рекурсивную (Recursive), прямую стратегию (Direct) и их объединение RecDir относятся к стратегиям с одним выходом, поскольку они моделируют данные как функцию с одним выходом (с несколькими входами).

Совместная стратегия, также называемая «множество входов и много выходов» (MIMO, S^{MIMO}) (доступна только для моделей нелинейного прогнозирования) предполагает обучение

одной многовыходной модели на истории наблюдений $[y_1, y_2, \dots, y_T]$:

$$(8) \quad M^{MIMO}: [y_{T+H}, \dots, y_{t+1}] = F(y_t, \dots, y_{t-d+1}) + w,$$

где $t \in \{d, \dots, T-H\}$, $F: R^d \rightarrow R^H$ является векторной функцией, w – вектор шума с необязательно диагональной ковариационной матрицей.

Прогноз формируется за один шаг:

$$(9) \quad S^{MIMO}: [\hat{y}_{T+H}, \dots, \hat{y}_{t+1}] = \hat{F}(y_T, \dots, y_{T-d+1}).$$

Введение стратегии MIMO мотивировано необходимостью избегать моделирования по стратегии с одним выходом, которое игнорирует существование стохастических зависимостей между будущими значениями, что влияет на точность прогноза. S^{MIMO} формируется на основе стохастической зависимости между прогнозируемыми значениями, характеризующей временной ряд. Она позволяет избежать условной независимости h прогнозов прямой стратегии, а также накопления ошибок, свойственного рекурсивной стратегии. Однако необходимость сохранения стохастических зависимостей за счет использования одной модели имеет недостаток, поскольку ограничивает все прогнозируемые горизонты одной и той же структурой модели. Это ограничение может снизить гибкость подхода к прогнозированию. Для решения этой проблемы предложена стратегия объединения MIMO и стратегии Direct – DIRMO. Также применение стратегии MIMO требует большего объема данных при обучении и подвержено проблеме переобучения.

Прогнозирование обнаружения структурных сдвигов в процессе и динамики развития процесса после того как в нем произошёл структурный сдвиг является важным элементом формирования многошагового прогноза. Решение этой задачи позволяет проанализировать динамику ряда и своевременно обнаружить наличие изменений (структурных сдвигов).

Поэтому важнейшей составной частью задачи прогнозирования процессов со структурными сдвигами является задача мониторинга – обнаружения разладки временного ряда, характеризующей момент времени, связанный с существенными изменениями свойств ряда. Еще больше возрастает роль мони-

торинга, если изменение динамики наблюдаемого процесса на горизонте прогнозирования обусловлено влиянием событий внешней среды.

В частности, механизмы формирования цен на товарных рынках, состав, сила влияния и направление других факторов, включенных в модель прогноза, могут изменяться на горизонте прогноза. В этих условиях модель перестает описывать динамику прогнозируемой цены, становится непригодной для прогноза и требует замены. Причинами изменений могут служить изменения внешней среды (политические или экономические события, изменения в планах развития масштабных проектов, пандемия, военно-политические конфликты и пр.) и внутренние механизмы функционирования рынка (конкуренция, ожидания относительно будущего участников рынка, используемые ими стратегии). Поэтому необходимо отслеживать происходящие изменения, своевременно изменять и корректировать модели посредством мониторинга, целями которого являются: обнаружение изменений временных рядов (структурных сдвигов), выявление причин, которые вызвали или могут вызвать возникновение структурных сдвигов.

Для обнаружения и идентификации типов изменений нестационарных процессов используются два широко распространенных подхода [33]:

- 1) флуктуационные тесты, модифицированные для применения в режиме мониторинга;
- 2) алгоритмы последовательного обнаружения и идентификации изменений, в основе которых лежит подход А. Вальда.

Алгоритмы мониторинга, основанные на флуктуационных тестах, представляют собой модификацию ретроспективного теста рекурсивных оценок момента изменения свойств, границы которого расширяются с ростом объема выборки. Эти тесты просты и удобны в применении.

В основе алгоритмов последовательного обнаружения лежит схема обработки поступающих наблюдений с использованием статистики отношения правдоподобия. В качестве критерия качества обнаружения используется минимизация запаздывания при ограничениях на частоту ложных обнаруже-

ний и вероятность ошибочной идентификации. Последовательные алгоритмы ранее использовались преимущественно в технических приложениях. Работы, исследующие эффективность их применения для экономических данных, появились сравнительно недавно.

Для прогнозирования нестационарных процессов используются различные методы: построение параметрических моделей, сингулярный спектральный анализ, декомпозиция по эмпирическим модам, машины опорных векторов, нейронные сети и др. [21, 23]. Однако высокий уровень неопределенности в динамике процессов (например для таких объектов, как товарные рынки) создает проблемы для применения только количественных моделей прогнозирования.

Чтобы уменьшить эту неопределенность, некоторые авторы [48, 52] исследовали различные стратегии многошагового прогнозирования (обзор стратегий приведен выше) и показали, что наиболее эффективными для многошагового прогнозирования являются стратегии с несколькими выходными данными. Применение этих стратегий предполагает сохранение стохастических зависимостей между членами прогнозируемой последовательности. Этим условиям удовлетворяют регулярные процессы, когда динамика ряда может быть с достаточной точностью описана некоторой динамической системой. Проблема, однако, в том, что динамика временных рядов часто меняется со временем, а причины изменения связаны с влиянием внешних событий. Поэтому в таких ситуациях более эффективен подход, основанный на извлечении и учете в прогнозе дополнительной информации о событиях прогнозируемого периода, влияющих на прогнозируемый процесс [58].

Однако эта информация не всегда может быть своевременно отражена в данных. Для нестационарных процессов, изменение динамики которых часто связано с внешними событиями, вызывающими структурные сдвиги, все больше внимания привлекают подходы, в основе которых лежит применение для формирования прогнозов информации о событиях, влияющих на значение прогнозируемого показателя на горизонте прогноза. Эта информация, извлекаемая из дополнительных источников,

далеко не всегда представляет собой временные ряды. Она может быть получена в результате:

- анализа текущей ситуации экспертами и аналитиками, прогнозистами и лицами, принимающими решения,
- поиска (по ключевым словам), анализа и обработки информации из разнородных информационных источников.

Поэтому возникает проблема как получить разнотипную информацию и как включать ее в прогноз.

Обобщая результаты анализа основных подходов и направлений количественного прогнозирования применительно к нестационарным процессам со структурными сдвигами на горизонте прогноза, можно сделать вывод, что существующие на сегодня методы количественного прогнозирования не обеспечивают требуемое качество прогноза в силу возрастающей неопределенности в динамике обозначенного объекта на горизонте среднесрочного прогнозирования. На краткосрочном горизонте все три группы методов, обозначенные в данном разделе, способны формировать прогнозы заданной точности; по мере расширения горизонта такая способность снижается (рис. 2).

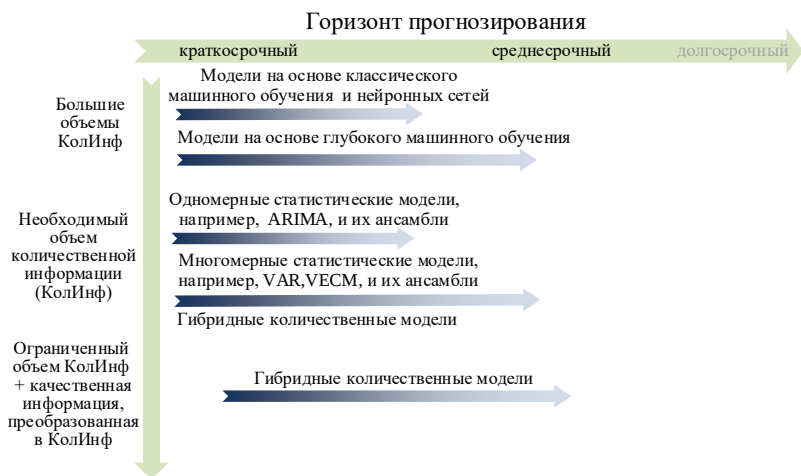


Рис. 2. Применимость количественных методов на горизонте прогнозирования

При этом многомерные статистические модели и их ансамбли, модели на основе глубокого обучения и количественные гибридные количественные модели можно рассматривать как основу для развития интеграционного прогнозирования.

Также важной характеристикой методов количественного прогнозирования является интерпретируемость (объяснимость) полученных результатов их применения для принятий решений (рис. 3).

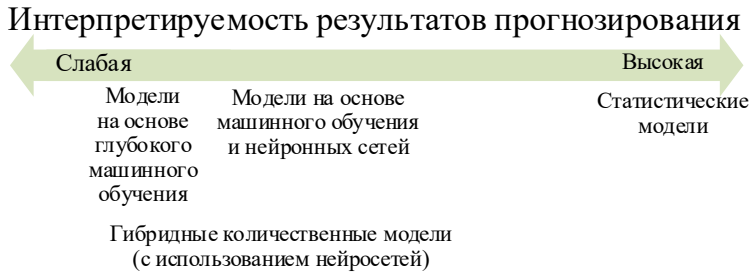


Рис. 3. Интерпретируемость (объяснимость) результатов применения методов для принятия решений

Еще одним важным показателем применимости методов является точность прогноза. Для ее оценки могут быть использованы различные меры оценки ошибки прогнозирования (MAE, MAPE, MSE, RNSE, MAP, R2, MAD и др.). Выбор зависит от специфики решаемой задачи, от используемых методов (в частности, когда прогноз формируется разными методами) и т.п. Статистика использования различных мер точности для разных областей здесь не рассматривается, это тема отдельного обзора.

4. Прогнозирование на основе интеграционных методов, использующих количественные данные, суждения и информацию из разнородных источников

Одно из важных преимуществ интеграции количественных методов и методов извлечения и обработки суждений и инфор-

мации из разнородных источников связано с их способностью обрабатывать различные типы информации.

Интеграционные методы – это методы, позволяющие учитывать разнотипную информацию (структурированную и неструктурированную) на разных этапах решения задачи, и тем самым прямым или косвенным способом включать ее в формируемый прогноз.

4.1. ТИПЫ ИНФОРМАЦИИ И ПОДХОДЫ К ЕЕ ИСПОЛЬЗОВАНИЮ НА РАЗЛИЧНЫХ ЭТАПАХ ПРОЦЕССА ПРОГНОЗИРОВАНИЯ

С увеличением горизонта прогноза возрастает потребность в учете различных типов информации для решения задач прогнозирования нестационарных процессов в условиях структурных сдвигов и повышения точности формируемых прогнозов.

При применении интеграционных методов можно выделить три типа информации (знаний), включаемых в прогноз [40, 43, 50].

Контекстная информация (contextual information) – конкретная информация, доступная в среде прогноза. Это информация, отличная от временных рядов и общего опыта, которая помогает в объяснении, интерпретации и прогнозировании поведения временных рядов. Такая информация (включающая как структурированные данные, так и неструктурированные текстовые данные) может собираться из различных информационных источников с использованием поисковых систем или специализированных программных средств интеллектуального анализа текстов.

Знания предметной области (domain knowledge) – знания, которые практикующие специалисты, эксперты получают благодаря опыту в рамках своей работы. Знание предметной области является результатом применения человеческой интерпретации к контекстной информации (например, информации о значимых событиях внешней среды, которые не отражены в данных текущих наблюдений процесса, но могут проявиться в них только с задержкой).

Технические знания (technical knowledge) – знания формальных методов и процедур анализа данных и прогнозирования.

Суждения, представленные контекстной и/или экспертной информацией, могут включаться на разных этапах процесса прогнозирования [40, 44]:

1) использование суждений при выборе и разработке количественного прогноза (ввод информации в статистический процесс путем выбора соответствующих моделей и параметров);

2) декомпозиция проблемы (задачи) – разбиение ее на части для формирования прогнозов по этим частям с последующим объединением полученных отдельных прогнозов, вместе определяющих прогноз целевого значения прогнозируемого объекта. Формы декомпозиции (мультипликативная или аддитивная) определяются в соответствии с характером проблемы и знаниями о взаимосвязи между ее компонентами [15, 32, 40];

3) комбинирование (совместное использование) качественных прогнозов (на основе суждений) и количественных (цифровых) прогнозов (распространенная линейная форма объединения представлена уравнением (1));

4) субъективная корректировка (judgmental adjustment) количественных прогнозов:

$$(10) \hat{F}_{fin} = \hat{F}_0 + \Delta\hat{F},$$

где $\hat{F}_0$ – статистический прогноз, полученный формальным методом; $\Delta\hat{F}$ – корректировка прогноза $\hat{F}_0$, полученная на основе экспертной информации.

Несмотря на то, что наиболее эффективный способ использования суждений – это ввод данных в статистический процесс путем выбора соответствующих моделей и параметров (подход 1), сделать это часто невозможно в условиях, когда прогнозист получает информацию несвоевременно [50]. К тому же роль суждения в этом подходе недостаточно изучена. Основное внимание уделялось автоматическим (статистическим) методам выбора моделей (в [45] данный подход получил название прогнозирование на основе признаков (feature)), несмотря на то что

специалисты по прогнозированию часто склонны игнорировать рекомендации системы и применять суждения при выборе модели прогнозирования. Поэтому вопрос о выборе подходящих моделей прогнозирования по-прежнему остается актуальным. Кроме того отмечается, что по мере увеличения объема информации специалисты не в состоянии обрабатывать ее эффективно и одновременно, поэтому декомпозицию (подход 2) целесообразно включать на этапе выбора модели [44].

В сообществе прогнозистов применяются два подхода с использованием признаков временных рядов: выбор модели на основе признаков и сочетание моделей на основе признаков. Обоснованием применения указанных подходов является то, что единственная модель не всегда работает наилучшим образом для всех временных рядов. Поэтому вместо выбора одной модели для всех данных предлагается использовать признаки для получения наиболее подходящей модели или оптимальной комбинации моделей-кандидатов. Признак временного ряда может быть любым статистическим представлением характеристик временного ряда. Эти признаки включают сводные данные временных рядов с точки зрения их корреляционной структуры, распределения, стационарности, свойств масштабирования и соответствия набору моделей временных рядов [29, 45].

Для выбора моделей (или их комбинаций) применяют автоматические (статистические) алгоритмы на основе информационных критериев [44] (таких как информационный критерий Акайке (AIC), байесовский информационный критерий (BIC) и т.д.) [22], на основе оценки эффективности различных моделей на наборе валидационных данных (перекрестная валидация) [42]. Наконец в рамках автоматического выбора модели оцениваются различные характеристики временных рядов (таких как тренд, сезонность, случайность, асимметрия, прерывистость, изменчивость, количество доступных наблюдений), а также учитываются переменные для принятия решений (таких как горизонт прогноза). С учетом этих характеристик (признаков) предлагается подход к выбору подходящих моделей на основе экспертных правил ([12, 13, 15, 24]) или с развитием машинного обучения – процедур мета-обучения процесса автоматического

получения знаний для выбора моделей (например, в исследовании [47] применяются деревья решений). Однако такие подходы очень чувствительны к выбранным правилам или функциям мета-обучения. Общепринятого набора таких правил не существует. Также предложен автоматизированный подход к извлечению признаков временных рядов на основе их визуализации, в котором признаки, извлеченные из изображений, используются для оценки оптимальных весов комбинации прогнозов [37].

Все перечисленные подходы для автоматического выбора моделей основаны на статистике или могут быть реализованы с помощью алгоритмического процесса. Общим для них является то, что выбор основан на исторических данных. Ни один из этих алгоритмов не может оценивать прогнозы, когда соответствующие фактические данные еще недоступны [44].

Подходы 3 («комбинирование») и 4 («субъективная корректировка количественных прогнозов») относятся к основным подходам к интеграции суждений со статистическими прогнозами [15, 43] и наиболее представлены в публикациях.

Комбинирование происходит, когда используется некоторая форма процедуры объединения статистических прогнозов и прогнозов на основе суждений [31, 43].

Подход 4 представляет собой апостериорное включение субъективной корректировки количественного прогноза с учетом дополнительной информации.

Субъективная корректировка полученных прогнозов может быть полезной при определенных условиях, когда высокая изменчивость временных рядов ухудшает эффективность статистического моделирования. Эта изменчивость вызвана структурными сдвигами, присущей им случайностью и неопределенностью, которые не могут быть объяснены причинными факторами, но могут быть отражены в контекстной информации (например, о событиях, которые не отражены в модели) [15]. Такая корректировка прогноза может быть в «ручном» режиме, проводимая прогнозистом на основе экспертных суждений, или с использованием формализованных процедур.

Например, в [40] представлен подход к структурированию и формализации экспертных знаний для коррекции статистиче-

ского прогноза из-за предстоящих особых событий. Структурирование достигается с помощью четырех формализованных факторов (transient factors, quantum jump factor, transferred impact factor, trend change factor), определяющих типы динамики, для идентификации событий и соотнесения их с этими факторами. Эксперт-прогнозист оценивает силу (веса) воздействия выявленных корректирующих факторов, ассоциированных с событиями, эти эффекты суммируются и добавляются к исходному статистическому прогнозу. Развитие данного подхода представлено в недавнем исследовании [19], где прогноз спроса на новую продукцию формируется на основе структурирования информации из социальных сетей и субъективной экспертной корректировки прогноза (с использованием корректирующих факторов) с учетом изменения настроений под воздействием событий, ассоциируемых с этими факторами.

Подход 4) относится к типичной практике в компаниях и представляет собой основную конкурирующую альтернативу комбинированному процессу интеграции суждений в прогнозы [15].

Как отмечалось в разделе 3, объединение (подход 3)) наиболее эффективно, когда объединенные прогнозы не коррелированы и привносят в процесс прогнозирования различные виды информации [14, 40]. При этом в [49] показано, что 1) контекстная информация более важна по сравнению с техническими знаниями при объединении прогнозов; 2) чем более нестабильны временные ряды, тем больше контекстной информации требуется для хорошей комбинации.

Интеграционное прогнозирование при объединении прогнозов часто использует ту же идею нахождения весов при комбинации количественных прогнозов (раздел 3), при этом допускается использование суждений. В [7] проведен сравнительный анализ методов формирования весов, использующих экспертную информацию, для объединения частных прогнозов в общий; субъективная корректировка (подход 4) также рассматривается в контексте объединения количественной и экспертной информации при формировании итогового прогноза.

Комбинированный подход (подход 3) к прогнозированию опирается на использование различных типов информации

и/или методов (алгоритмов). В последнее время все большее распространение получают гибридные модели, применение которых подразумевает обработку разных типов информации несколькими моделями для формирования конечного прогноза.

Одним из таких подходов является улучшение прогнозной модели целевого показателя за счет включения в нее неструктурированных данных, обработанных алгоритмами поисковых систем и используемых в качестве объясняющих переменных изменения динамики целевого показателя [30].

В частности, для учета экологических событий в прогнозе авторы [54] предложили методику многомасштабного прогнозирования цен на нефть, используя многофакторные данные поисковых систем в разных временных масштабах.

В [30] представлен обзор применения традиционных статистических моделей и методов машинного обучения для прогнозирования туристического спроса с использованием неструктурированной информации. Отнесение данного подхода к интеграционным методам обусловлено тем, что несмотря на совершенствование поисковых алгоритмов и интеллектуального анализа текстов, остается значимой проблема качественного отбора ключевых слов, являющегося наиболее важным этапом для формирования поисковых запросов в массивах неструктурированной информации. Общепринятого метода, позволяющего систематическим образом осуществлять такой отбор, не существует [30]; как правило, полагаются на экспертное знание предметной области и использование различных терминов, связанных с прогнозируемым процессом.

Перспективным направлением решения задач прогнозирования временных рядов является когнитивный анализ и моделирование ситуаций, о чем свидетельствуют недавние обзоры [26, 41].

В основе этого направления лежит применение нечетких когнитивных карт¹ (НKK) – формализованных моделей пред-

¹ В зарубежных публикациях общепринят термин «нечеткая когнитивная карта» (НKK) (предложен Б. Коско [36]) при исследовании динамических систем в контексте мягких вычислений. НKK представляют собой графовую

ставления знаний о причинно-следственных влияниях между значимыми факторами (концептами) ситуации.

Можно выделить два направления развития методов на основе НКК в этой области.

Одно из них ориентировано на преобразование временного ряда конкретного параметра (или нескольких параметров) в НКК для последующего формирования прогноза динамики этого параметра (параметров) (см., например, [51]) с использованием различных методов обучения.

В контексте целей настоящего обзора представляет интерес второе направление по развитию интеграционных методов – гибридный подход, где НКК используются совместно со статистическими моделями или нейронными сетями. В этом направлении можно выделить ряд исследований, где формирование и/или коррекция прогнозной модели опирается на данные о событиях (инфоповодах) внешней среды, извлеченных из разнородных источников информации и обработанных с использованием НКК (см. [9, 16, 35]).

В [35] представлена гибридная модель – знаковая НКК и нейронная сеть – для прогнозирования рядов финансовых данных. В знаковой НКК отражаются только направления положительного (усиливающего) или отрицательного (ослабляющего) причинно-следственного влияния между факторами (концептами). Гибридная модель составляет основу автоматизированной системы KBNMiner, включающей базу экспертных знаний, формируемую по НКК различных предметных областей; информационно-поисковую подсистему для автоматического сбора новостной информации о событиях из интернета, где направленный поиск организован с использованием концеп-

структуру представления причинно-следственных рассуждений, где нечеткость обусловлена размытой (неоднозначной) степенью причинности между размытыми причинными концептами (факторами) [36] (как правило, степень причинности (сила влияния) между концептами лежит в диапазоне $[-1, 1]$ (или $[0, 1]$)). В отечественных публикациях большее распространение получил термин «когнитивная карта» при том же смысле. В данном обзоре используется НКК независимо от употребляется термина в цитируемых публикациях.

тов НКК; подсистему обработки выделенных положительных и отрицательных событий на основе анализа структуры причинно-следственных влияний НКК для последующего ввода результатов в прогнозную модель нейронной сети в виде параметра, характеризующего относительную силу (положительного или отрицательного) воздействия на прогнозируемый показатель, вместе с другими связанными с ним финансовыми показателями. Эмпирическая проверка на задаче прогнозирования процентных ставок подтвердила, что учет качественной информации оказывает значительное влияние на эффективность нейронной сети при прогнозировании.

В [2, 16] рассматривается задача прогнозирования цен на сырьевые товары на год вперед с разбивкой по месяцам. Сложность решения задачи обусловлена ростом уровня неопределенности прогноза по мере увеличения горизонта прогнозирования. Для уменьшения неопределенности предложена гибридную модель формирования и корректировки ежемесячного прогноза цен на год вперед, которая основана на использовании ансамблей многомерных цифровых моделей временных рядов, построенных в различных временных масштабах. Помимо прогнозируемых показателей и оказывающих влияние на прогноз временных рядов цен и макропоказателей модель использует экспертно-аналитическую информацию, обработанную с помощью НКК ситуации. Для коррекции прогнозной модели на горизонте прогнозирования применяется совместный мониторинг цен на товарных рынках [17], включающий 1) цифровой мониторинг для выявления структурных сдвигов во временном ряду прогнозируемого процесса (и рядах связанных с ним процессов); 2) ситуационный мониторинг внешней среды (по разнородным информационным источникам) и формирование сигналов по результатам сценарного моделирования на НКК возможных последствий влияния значимых внешних событий (инфоповодов) на изменение прогнозируемого процесса и связанных с ним процессов; 3) управление обменом информацией между 1) и 2) и формирование выходных агрегированных сигналов. Работоспособность предложенного подхода подтверждена на примере прогнозирования цен на черный лом на 2019 г.

В [3, 9] предложена модель модулярной системы прогнозирования временных рядов, включающая три основных модуля. Гибридная нейро-нечеткая сеть формирует прогноз временного ряда – количественный прогноз, результаты которого верифицируются по критерию заданной точности. НКК, учитывающая факторы причинно-следственного влияния на прогнозируемый показатель на основе собранных данных о событийном влиянии на временной ряд, формирует прогноз с оценкой влияния события на прогнозируемый показатель. Итоговый прогноз формируется моделью нейронной сети, агрегирующей результаты работы первых двух моделей.

В контексте повышения точности прогнозов также рассматривается возможности сочетания подходов 3) и 4). В [28] представлен подход к комбинированию частных прогнозов, скорректированных экспертами; при этом показано, что комбинированные прогнозы, основанные на модели, могут быть превзойдены комбинированными прогнозами, полученными на основе частных прогнозов, скорректированных экспертами, даже когда сами скорректированные частные прогнозы менее точны, чем прогнозы, лежащие в основе модели.

Таким образом, для решения задачи прогнозирования нестационарных процессов в условиях структурных сдвигов на длительном горизонте прогнозирования H необходимо учитывать различные типы информации:

$$(5) \quad \hat{F}_{fin} = f(\{Y_T^1, \dots, Y_T^k, \dots, Y_T^N\}, H, I_{cont}, I_{exp}, I_{tech}),$$

где $\{Y_T^1, \dots, Y_T^k, \dots, Y_T^N\}$ – количественные данные в виде многомерных рядов; I_{cont} – контекстная информация; I_{exp} – экспертная информация; I_{tech} – техническая информация.

Кроме того, специфика задачи требует применения в качестве методологической базы ее решения – стратегии многошагового (многомасштабного) прогнозирования многомерных рядов.

Обобщая проведенный анализ основных направлений и подходов к интеграционному прогнозированию нестационарных процессов со структурными сдвигами (раздел 3 и раздел 4.1), на рис. 4 систематизированы модели и методы, приме-

нимые для решения задачи среднесрочного прогнозирования таких процессов. Слева приведена стандартная схема решения указанной задачи, справа – модели и методы, которые могут комбинироваться с учетом длины горизонта, типа информации и данных о прогнозируемом процессе, их доступности и полноте; требований к точности прогноза.



Рис. 4. Основные модели и методы, применимые для решения задачи среднесрочного прогнозирования нестационарного процесса со структурными сдвигами

4.2. ОСНОВНЫЕ ПРОБЛЕМЫ РАЗРАБОТКИ И ПРИМЕНЕНИЯ ИНТЕГРАЦИОННЫХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ И ПУТИ ИХ РЕШЕНИЯ

Несмотря на признанную значимость интеграционных методов прогнозирования с включением суждений и неструктурированной информации, основными проблемами их разработки и применения в контексте повышения точности формируемых прогнозов являются предвзятость специалистов по прогнозированию, экспертов и аналитиков, недостаточная технологичность методов структурирования неоднородной информации, необхо-

димых для выделения и формализации полезной информации, формирования и обработки сигналов на ее основе для включения в прогноз [14, 40, 50].

Сложность своевременного получения полезной и достоверной информации о характере и тенденциях изменения процессов воздействия внешней среды на объект прогнозирования обусловлена, с одной стороны, недостатком информации и понимания существенных разнородных факторов – движущих сил, определяющих изменение этих процессов, а с другой стороны – проблемой большого «шума» в данных, характеризующих исследуемые процессы, из-за их избыточной доступности. Такая доступность обусловлена большим количеством источников и порождаемыми ими большими потоками информации в условиях повсеместного внедрения цифровых и информационно-телекоммуникационных технологий, что может привести к информационной перегрузке, влияющей на способность специалистов определять достоверность, актуальность и надежность имеющейся информации [57]. К тому же обработка в таких объемах и не требуется прогнозисту для решения поставленных задач. Кроме того, контекстная информация часто представляется в неколичественной форме, и ее количественная оценка может быть сложной задачей. Поэтому такую информацию сложно интегрировать с данными временных рядов.

Путь решения обозначенных проблем интеграционного прогнозирования специалисты и исследователи в области анализа и прогнозирования временных рядов в контексте повышения точности прогнозов видят в совершенствовании систем поддержки прогнозирования (СПП), предназначенных не для замены прогнозистов, а для структурированной и систематической поддержки деятельности прогнозистов, экспертов и аналитиков в жизненном цикле решения практических задач прогнозирования [10, 11, 15, 27].

Согласно [15] СПП определяется как программное обеспечение, фреймворк (готовый набор инструментальных средств) и/или структурированная процедура, которые могут использовать экспертные суждения, разнотипную информацию и количественные методы для интеграции исторических данных

и контекстной информации с целью оказания значимой поддержки прогнозисту для формирования более точных прогнозов и анализа результатов.

На практике СПП не получили широкого распространения. Большинство компаний используют коммерческие пакеты для прогнозирования спроса на продукцию. Прогнозы, создаваемые этими пакетами, часто корректируются на основе управленческих суждений. Причинами медленного развития и внедрения практических СПП являются: высокая стоимость разработки и внедрения СПП; сложность обслуживания и модернизации СПП в условиях быстрых изменений в типе и объеме данных и информации. Еще одна проблема связана с тем, что существующие СПП систематически не собирают и не включают экспертную информацию в прогнозы, что приводит к системному недоверию и ручной корректировке результатов прогнозистами [15].

Так как суждения играют важную роль в жизненном цикле решения практических задач прогнозирования, начиная от выбора моделей и до коррекции прогноза, их невозможно исключить из этого процесса даже при наличии хорошо продуманной СПП [43, 44]. На рис. 5 представлена схема расширенного информационного обеспечения основных этапов прогнозирования нестационарных процессов в условиях структурных сдвигов (в пунктирном прямоугольнике выделен необязательный этап – декомпозиция задачи).

Как видно на рис. 5, экспертные знания являются основой информационного обеспечения, так как используются на всех этапах жизненного цикла прогнозирования, но следует отметить, что периодичность их актуализации на разных этапах различна.

СПП может рассматриваться как способ интеграции моделей и методов прогнозирования с учетом разнотипной информации, включаемой в прогноз.

Вопросы разработки принципов и требований к формированию функционала такой СПП в контексте решения задачи прогнозирования нестационарных процессов в условиях структурных сдвигов не рассматриваются в данном обзоре, так как

это требует серьезной проработки и не является предметом данного обзора.



Рис. 5. Схема расширенного информационного обеспечения основных этапов решения задачи прогнозирования нестационарных процессов в условиях структурных сдвигов

Однако можно выделить ряд общих взаимосвязанных принципов ее построения для решения указанных задач в контексте повышения точности прогнозов на заданном горизонте прогнозирования.

Соответствие функционала сложности решаемых задач, адаптивность СПП. Понимание характеристик задач и соответствующих потребностей в поддержке позволяет создавать СПП, которые лучше соответствуют потребностям пользователей и адаптируются к ним.

Технологичность процесса решения задачи прогнозирования, обеспечивающая систематическую и структурированную процедуру решения задачи. К тому же подгонка технологической поддержки к характеристикам и особенностям задач спо-

способствует выявлению пробелов между функциональностью системы и потребностями пользователей.

Поддержка многомодельности (гибридный подход) в СПП, дающая возможность работы со структурированной и неструктурированной информацией (рис. 5) в жизненном цикле решения задачи прогнозирования. Для облегчения выбора моделей – поддержка теоретическими знаниями по свойствам и поведению моделей; накопление базы знаний по положительному и отрицательному опыту применения моделей.

Включение модели представления знаний предметной области с целью повышения качества направленного поиска и фильтрации неструктурированной информации при выделении полезной контекстной информации для формируемого прогноза. Например, в качестве моделей представления знаний можно использовать НКК, графы знаний.

Снижение когнитивной нагрузки и ресурсоемкости обработки информации человеком. СПП, которые дополняют возможности человека по обработке информации, могут повысить эффективность решений, поскольку они снижают когнитивную нагрузку на человека при решении сложных задач среднесрочного и долгосрочного прогнозирования.

Информативность, позволяющая обеспечить постоянный анализ результатов в жизненном цикле решения задачи.

5. Заключение

Несмотря на многолетнюю историю исследований в области прогнозирования с использованием суждений (Forecasting with Judgment), публикуемых зарубежных обзорах, затрагивающих различные аспекты этого направления, актуальность его развития сохраняется, особенно для решения задач среднесрочного и долгосрочного прогнозирования процессов в сложных динамических системах. В этих системах протекают нестационарные процессы, динамика которых может изменяться не только вследствие эндогенных изменений, но и под воздействием экзогенных факторов в окружающей среде объекта на длительном горизонте прогнозирования. В таких условиях повыше-

ние точности прогноза зависит не только от используемых количественных методов прогнозирования, но и от методов, способных обнаруживать и учитывать информацию о будущих событиях, изменяющих динамику целевого показателя в интервале между прогнозами, отсутствующей в данных на момент составления прогноза.

Обобщая результаты проведенного обзора, можно сделать вывод, что для задач прогнозирования нестационарных процессов в условиях структурных сдвигов на длительном горизонте прогнозирования является перспективным применение интеграционных методов. Такие методы должны сочетать в себе возможности гибридных моделей, комбинирования современных количественных методов и методов обработки всей доступной полезной контекстной и экспертной информации об объекте прогнозирования, учитываемой на разных этапах решения задачи прогнозирования.

При этом, учитывая сложность решения задач среднесрочного и долгосрочного прогнозирования для таких объектов, необходимость решения проблемы полноты и достаточности включаемой разнородной информации в формируемый прогноз, развитие таких методов перспективно по пути их интеграции с СПП.

Литература

1. АВДЕЕВА З.К., КОВРИГА С.В., ГРЕБЕНЮК Е.А. *Анализ методов среднесрочного прогнозирования процессов со структурными сдвигами на финансовых и товарных рынках* // Проблемы управления. – 2024. – №5 (в печати).
2. АВДЕЕВА З.К., КОВРИГА С.В., ГРЕБЕНЮК Е.А. *Формирование среднесрочных месячных прогнозов цен на сырье на основе экспертной и количественной информации* // Автоматизация в промышленности. – 2022. – №5. – С. 38–45.
3. АВЕРКИН А.Н., ЯРУШЕВ С.А., ПАВЛОВ В.Ю. *Когнитивные гибридные системы поддержки принятия решений и прогнозирования* // Программные продукты и системы. – 2017. – №4. – С. 632–642.

4. АНДРИАНОВА Е.Г., ГОЛОВИН С.А., ЗЫКОВ С.В. и др. *Обзор современных моделей и методов анализа временных рядов динамики процессов в социальных, экономических и социотехнических системах* // Russian Technological Journal. – 2020. – №8(4). – С. 7–45. – DOI: <https://doi.org/10.32362/2500-316X-2020-8-4-7-45>.
5. АСТАФЬЕВА Е.В., ТУРУНЦЕВА М.Ю. *Анализ возможностей улучшения качества прогнозов цен на природные ресурсы методами комбинирования на основе регрессионных оценок весов* // Экономическое развитие России. – 2023. – Т. 30, №12. – С. 24–33.
6. СЕМЕНОВ А.В. *Анализ временных рядов в политической науке: возможности и ограничения* // Политическая наука. – 2021. – №1. – С. 76–97.
7. СУРКОВ А.А. *Объединение экономических прогнозов с использованием экспертной информации* // Статистика и экономика. – 2019. – №16(5). – С. 4–14.
8. ФРЕНКЕЛЬ А.А., СУРКОВ А.А. *Методологические подходы к улучшению точности прогнозирования путем объединения прогнозов* // Вопросы статистики. – 2015. – №8. – С. 17–36.
9. ЯРУШЕВ С.А., АВЕРКИН А.Н., ФЕДОТОВА А.В. *Модулярная модель прогнозирования временных рядов на основе нейро-нечетких сетей и когнитивного моделирования* // Нечеткие системы и мягкие вычисления. – 2017. – Т. 12, вып. 2. – С. 159–168.
10. ABOLGHASEMI M., HURLEY J., ESHRAGH A. et al. *Demand forecasting in the presence of systematic events: Cases in capturing sales promotions* // Int. Journal of Production Economics. – 2020. – Vol. 230. – 107892.
11. ADYA M., LUSK E.J. *Designing Effective Forecasting Decision Support Systems: Aligning Task Complexity and Technology Support* / C. Jao (ed.) Decision Support Systems. – IntechOpen Limited. United Kingdom, 2012.
12. ADYA M., LUSK E.J. *Development and validation of a rule-based time series complexity scoring technique to support design of adaptive forecasting DSS* // Decision Support Systems. – 2016. – No. 83(3). – P. 70–82

13. ARINZE B. *Selecting appropriate forecasting models using rule induction* // Omega. – 1994. – No. 22(6). – P. 647–658.
14. ARMSTRONG J.S. *Combining forecasts* // In: Principles of Forecasting: A Handbook for Researchers and Practitioners / J.S. Armstrong (ed.). – Springer Nature, 2001. – P. 417–440.
15. ARVAN M., FAHIMNIA B., REISI M. et al. *Integrating human judgement into quantitative forecasting methods: A review* // Omega. – 2019. – Vol. 86. – P. 237–252. – DOI: <https://doi.org/10.1016/j.omega.2018.07.012>.
16. AVDEEVA Z.K., GREBENYUK E.A., KOVRIGA S.V. *Raw Material Price Forecasting on Commodity Markets: Application of Expert and Quantitative Information* // Advances in Systems Science and Applications. – 2022. – Vol 22, No. 4. – P. 126–143.
17. AVDEEVA Z.K., GREBENYUK E.A., KOVRIGA S.V. *Detection of structural shifts in commodity markets in the mode of situation and digital monitoring* // IFAC-PapersOnLine. – 2023. – Vol. 56, Iss. 2. – P. 7778–7783.
18. AZMY W.M., ATIYA A.F., EL-SHISHINY H. *Forecast Combination Strategies for Handling Structural Breaks for Time Series Forecasting* // In: Multiple Classifier Systems. MCS 2010. Lecture Notes in Computer Science / Eds.: N.E. Gayar, J. Kittler, F. Roli. – Springer, Berlin, Heidelberg. – 2010. – Vol. 5997. – DOI: https://doi.org/10.1007/978-3-642-12127-2_25.
19. BADULESCU Y., CAÑAS F., CHEIKHROUHOU N. *Judgmental adjustment of demand forecasting models using social media data and sentiment analysis within industry 5.0 ecosystems* // Int. Journal of Information Management Data Insights. – 2024. – Vol. 4, Iss. 2. – 100272.
20. BATES J.M., GRANGER C.W.J. *The combination of forecasts* // Journal of the Operational Research Society. – 1969. – No. 20(4). – P. 451–468.
21. BAYÜKŞAHINA Ü., ERTEKIN S. *Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition* // Neurocomputing. – 2019. – Vol. 361. – P. 151–163.

22. BRUNHAM K.P., ANDERSON D.R. *Model selection and multi-model inference: A practical information-theoretic approach*. – Springer-Verlag, Heidelberg, 2002. – 490 p.
23. CHENG C., SA-NGASOONGSONG A., BEYCA O. et al. *Time series forecasting for nonlinear and non-stationary processes: a review and comparative study* // IISE Trans. – 2015. – Vol. 47(10). – P. 1053–1071.
24. COLLOPY F., ARMSTRONG J.S. *Rule-Based Forecasting: Development and Validation of an Expert Systems Approach to Combining Time Series Extrapolations* // Management Science. – 1992. – No. 38(10). – P. 1394–1414.
25. DE STEFANI J., BONTEMPI G. *Factor-Based Framework for Multivariate and Multi-step-ahead Forecasting of Large-Scale Time Series* // Front. Big Data. – 2021. – Vol. 4. – 690267.
26. FELIX G., NÁPOLES G., FALCON R. et al. *A review on methods and software for fuzzy cognitive maps* // Artif Intell Rev. – 2019. – Vol. 52. – P. 1707–1737.
27. FILDES R., GOODWIN P., LAWRENCE M. *The design features of forecasting support systems and their effectiveness* // Decision Support Systems. – 2006. – Vol. 42, Iss. 1. – P. 351–361.
28. FRANCES F.H., VAN DIJK D. *Combining expert adjusted forecasts* // Journal of Forecasting. – 2019. – Vol. 38, Iss. 5. – P. 415–421.
29. FULCHER B.D., JONES N.S. *Highly comparative feature-based time-series classification* // IEEE Trans. on Knowledge and Data Engineering. – 2014. – No. 26(12). – P. 3026–3037.
30. GEORGAKIS A., PROFILIDIS V., BOTZORIS G.N. *Forecasting tourism demand using search engine data* // Proc. of the 10th Int. Congress on Transportation Research (ICTR2021). – Rhodes, Greece. – 2021.
31. GOODWIN P. *Integrating management judgment and statistical methods to improve short-term forecasts* // Omega. – 2002. – Vol. 30, Iss. 2. – P. 127–135.
32. GOODWIN P., FILDES R. *Forecasting with Judgment* // In: The Palgrave Handbook of Operations Research / Eds.: S. Salhi, J. Boylan. – Palgrave Macmillan, Cham, 2022. – DOI: https://doi.org/10.1007/978-3-030-96935-6_16.

33. GREBENYUK E.A. *Algorithms for Detecting Changes in the Properties of Non-Stationary Processes in the Monitoring Mode* // Proc. of the 14th Int. Conf. "Management of Large-Scale System Development" (MLSD). – Moscow: IEEE. – 2021. – URL: <https://ieeexplore.ieee.org/document/9600109>.
34. HOGARTH R.M., MAKRIDAKIS S. *Forecasting and planning: an evaluation* // Management science. – 1981. – No. 27(2). – P. 115–138.
35. HONG T., HAN I. *Integrated approach of cognitive maps and neural networks using qualitative information on the World Wide Web: the KBNMiner* // Expert Systems. – 2004. – No. 21(5). – P. 243–252.
36. KOSKO B. *Fuzzy cognitive maps* // Int J Man Mach Stud. – 1986. – No. 24(1). – P. 65–75.
37. LI X., KANG Y., LI F. *Forecasting with time series imaging* // Expert System with Applications. – 2020. – Vol. 160. – 113680.
38. LIVIERIS I.E., PINTELAS P. *A novel multi-step forecasting strategy for enhancing deep learning models' performance* // Neural Computing and Applications. – 2022. – Vol. 34. – P. 19453–19470.
39. MAKRIDAKIS S., HOGARTH R., GAB A. *Forecasting and uncertainty in the economic and business world* // Int. Journal of Forecasting. – 2009. – Vol. 25. – P. 794–812.
40. MARMIER F., CHEIKHROUHOU N. *Structuring and integrating human knowledge in demand forecasting: a judgmental adjustment approach* // Production Planning & Control. – 2010. – No. 21(4). – P. 399–412.
41. ORANG O., DE LIMA E SILVA P.C., GUIMARÃES F.G. *Time series forecasting using fuzzy cognitive maps: A survey* // Artif. Intell. Rev. – 2022. – DOI: <https://doi.org/10.1007/s10462-022-10319-w>.
42. ORD J.K., FILDES R., KOURENTZES N. *Principles of Business Forecasting*. – Wessex Press Publishing Co, 2017.
43. PERERA H.N., HURLEY J., FAHIMNIA B. et al. *The human factor in supply chain forecasting: A systematic review* // European Journal of Operational Research. – 2019. – Vol. 274(2). – P. 574–600.

44. PETROPOULOS F., KOURENTZES N., NIKOLOPOULOS K. et al. *Judgmental selection of forecasting models* // Journal of Operations Management. – 2018. – No. 60. – P. 34–46.
45. PETROPOULOS F., APILETTI D. et al. *Forecasting: theory and practice* // Int. Journal of Forecasting. – 2022. – Vol. 38(3). – P. 705–871.
46. POLAT C. *Forecasting as a strategic decision-making tool: A review and discussion with emphasis on marketing management* // European Journal of Scientific Research. – 2008. – Vol. 20(2). – P. 419–442.
47. PRUDÊNCIO R.B., LUDERMIR T.B. *Meta-learning approaches to selecting time series models* // Neurocomputing. – 2004. – No. 61. – P.121–137.
48. SAHOO D., SOOD N., RANI U. et al. *Comparative Analysis of Multi-Step Time-Series Forecasting for Network Load Dataset* // Proc. of the 11th Int. Conf. on Computing, Communication and Networking Technologies. – Kharagpur, India. – 2020. – P. 1–7. – DOI: <https://doi.org/10.1109/ICCCNT49239.2020.9225449>.
49. SANDERS N.R., RITZMAN L.P. *Bringing judgment into combination forecasts* // J. Oper. Manag. – 1995. – No. 13. – P. 311–321.
50. SANDERS N.R., RITZMAN L.P. *Judgmental Adjustment of Statistical Forecasts* // In: Principles of Forecasting: A Handbook for Researchers and Practitioners / J.S. Armstrong (ed.). – Springer Nature, 2001. – P. 405–416.
51. SHAN D., LU W., YANG J. *The data-driven fuzzy cognitive map model and its application to prediction of time series* // Int. Journal of Innovative Computing, Information and Control. – 2018. – No. 14(5). – P. 1583–1602.
52. TAIEB S., BONTEMPI G., ATIYA A. et al. *A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition* // Expert System with Applications. – 2012. – Vol. 39(8). – P. 7067–7083. – DOI: <https://doi.org/10.1016/j.eswa.2012.01.039>.
53. TAIEB S.B., ATIYA A.F. *A bias and variance analysis for multistep-ahead time series forecasting* // IEEE Trans. Neural Netw. Learn. Syst. – 2016. – No. 27. – P. 62–76.

54. TANG L., ZHANG C., LI L. et al. *A multi-scale method for forecasting oil price with multi-factor search engine data* // Applied Energy. – 2020. – Vol. 257. – 114033. – DOI: <https://doi.org/10.1016/j.apenergy.2019.114033>.
55. THOMSON M.E., POLLOCK A.C., ÖNKAL D. et al. *Combining forecasts: Performance and coherence* // Int. Journal of Forecasting. – 2019. – Vol. 35, Iss. 2. – P. 474–484.
56. WANG X., HYNDMAN R.J., LI F. et al. *Forecast combinations: An over 50-year review* // Int. Journal of Forecasting. – 2023. – Vol. 39, Iss. 4. – P. 1518–1547. – DOI: <https://doi.org/10.1016/j.ijforecast.2022.11.005>.
57. WEBBY R., O'CONNOR M., LAWRENCE M. *Time-series forecasting using domain knowledge* // In: Principles of Forecasting: A Handbook for Researchers and Practitioners / J.S. Armstrong (ed.). – Springer Nature, 2001. – P. 389–403.
58. WEBBY R., O'CONNOR M., EDMUNDSON B. *Forecasting support systems for the incorporation of event information: an empirical investigation* // Int. J. Forecast. – 2005. – Vol. 21. – P. 411–423.

INTEGRATING FORECASTING OF NON-STATIONARY PROCESSES REPRESENTED BY TIME SERIES. OVERVIEW

Zinaida Avdeeva, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand. Sc., Senior Researcher (avdeeva@ipu.ru).

Svetlana Kovriga, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Researcher (kovriga@ipu.ru).

Abstract: The review covers the main directions and approaches to integration forecasting of non-stationary processes represented by time series. The key source of non-stationarity generation is rapid and poorly predictable changes in the external environment, under the influence of which structural shifts occur in the complex economic and socio-political processes. The solution of problems of forecasting the dynamics of such objects in the context of improving the accuracy of the forecast becomes more complex as the forecast horizon increases. It determines the need for models and methods capable of processing heterogeneous information. Integration methods are methods that allow, along with quantitative data, to take into account judgments (of forecasters, experts, analysts) and information from heterogeneous

information sources at different stages of problem solving, and thus directly or indirectly include them in the forecast being formed. The development of such methods is aimed at increasing the accuracy of the forecast through the use of all available information about the forecasting object, including data on endogenic and exogenic factors of influence on it. The review focused on the current state of integration forecasting, on the existing problems and ways to solve them.

Keywords: non-stationary processes, time series, forecasting, judgments.

УДК 338.27

ББК 65.054

*Статья представлена к публикации
членом редакционной коллегии А.В. Горбуновой.*

Поступила в редакцию 01.10.2024.

Опубликована 30.11.2024.