

Аннотация

Работа посвящена поиску методов прогноза преступности, основанных на универсальных закономерностях поведения сложных нелинейных иерархических систем, плохо поддающихся обычному моделированию. На примерах недельных данных по преступности г. Ярославля изучается динамика преступлений и предлагаются алгоритмы прогноза скачков тяжких преступлений, прогнозирующих порядка 70% скачков в 30% тревожного времени при 4% числа ложных тревог.

A serious crime flash-up forecast based on hierarchy of criminality behavior

Abstract

This work is devoted to a quest of criminality forecast methods based on universal behavior rules of complex nonlinear hierarchical systems which are poorly amenable to routine modeling. Using a weekly criminality data from Yaroslavl for 1993-2001 years we study crimes' dynamics and propose serious crime flash-up forecast algorithms forecasting about 70% of flash-ups during 30% of alarm time with 4% of false alarms.

ВВЕДЕНИЕ

В своей жизни человек постоянно имеет дело с прогнозами, зачастую не отдавая себе в этом отчета. Угадывание «правильной» кассы в универсаме, правильное предвидение реакции собеседника, выбор одежды «по погоде» и многое другое является повседневной проверкой правильности выбранной данной личностью модели той или иной сферы жизни. Обладающие различной достоверностью (качеством) прогнозы погоды, курса валют, результатов выборов сопутствуют нам постоянно и могут не только помогать, но и мешать нам строить адекватно свою линию поведения.

Среди ученых отношение ко многим прогнозам, в особенности при отсутствии прозрачной модели процесса, достаточно скептическое. Напомним, однако, что отсутствие понимания физического механизма работы закона всемирного тяготения со времен Ньютона не препятствовало его успешному применению. Любым прогнозам очень далеко до точности и стройности закона всемирного тяготения. Тем не менее, в прогнозе, как и при обосновании научной гипотезы, совпадение предполагаемых результатов с реальной действительностью является единственным способом проверки адекватности используемого подхода.

Особенно важное место занимают прогнозы катастроф, кризисов и иных неожиданных резких изменений и редких явлений (засухи, наводнения, и т.д.). Человечество, в силу жизненной необходимости, всегда старалось предвидеть такие явления. Наибольшие успехи при прогнозе были достигнуты в медицине и метеорологии благодаря, как правило, хорошей статистике объектов прогноза и хорошо развитой системе мониторинга состояния объекта в реальном времени. Однако вполне успешными, даже и в этих областях, оказываются, как правило, лишь достаточно краткосрочные прогнозы. При переходе к большим временным интервалам на поведение исследуемой системы могут влиять дополнительные и пока еще плохо изученные факторы, а также принципиальная неустойчивость поведения большинства динамических систем. Примером тому является достаточно невысокая эффективность долгосрочных прогнозов погоды и непрекращающиеся дебаты о характере изменения климата.

В принципе существует два подхода к прогнозу поведения систем:

Первый предполагает разработку некоторой, достаточно точной модели исследуемой системы и получение прогноза на основе этой модели. Хорошим примером такого подхода являются прогнозы погоды, основанные на моделях тепломассопереноса и достаточно эффективные при прогнозе в малой окрестности рассматриваемой точки фазового пространства системы. Усложнение моделей и включение в них новых факторов могут позволить переходить к более точным долгосрочным прогнозам. Однако построение более точной модели требует детального описания исследуемого процесса, что для многих систем (на настоящий момент времени, а в ряде случаев и по принципиальным соображениям) нереально.

Весьма существенно также, что многие системы обладают внутренней неустойчивостью, способностью к резким изменениям своего состояния при малом изменении управляющих параметров. Такие изменения часто называются кризисами, катастрофами или чрезвычайными ситуациями и с математической точки зрения рассматриваются как особые точки фазового пространства. Возникновение таких неожиданных и резких изменений и служит основной причиной крупных ошибок в прогнозах, конструируемых в соответствии с первым подходом.

Второй путь нацелен преимущественно на прогноз именно резких изменений состояния систем и основан на использовании результатов синергетики – науки, исследующей универсальные закономерности поведения сложных нелинейных динамических систем, плохо поддающихся обычному моделированию. Подобные системы и их отдельные компоненты удобно рассматривать в рамках иерархической нелинейной модели. Одним из важных свойств таких систем является способность к самоорганизации, время от времени приводящей к резким изменениям поведения системы при незначительных изменениях внешних условий.

Одним из наиболее исследованных примеров большой нелинейной иерархической системы является литосфера Земли, представляющая собой систему блоков, разделенных разломами. Каждый из этих блоков делится на более мелкие – от тектонических плит протяженностью в тысячи километров до зерен горных пород миллиметрового размера и меньше. Тепловая конвекция Земли заставляет плиты двигаться и взаимодействовать между собой. Возникающие на разных масштабных уровнях напряжения могут приводить к разрушению соответствующих по размеру блоков. Разрушения мелких неоднородностей горной породы и микроблоков отвечают сейсмическому шуму, разрушение достаточно больших блоков фиксируются как землетрясения. Накапливающиеся в результате взаимодействия блоков напряжения имеют тенденцию распространяться от нижних уровней иерархии к верхним – обратный каскад, приводя к таким событиям, как сильные землетрясения. Разрядка напряжения землетрясениями может рассматриваться как перераспределение напряжений сверху вниз по иерархии – прямой каскад.

Другой пример иерархической системы – электорат, делящийся на блоки (страты) со своими групповыми интересами. Каждый из блоков делится на более мелкие группы, и т.д. вплоть до отдельного избирателя. Очевидно, что любое достаточно большое человеческое сообщество, как и соответствующие ему административные или общественные структуры, являются примерами подобных иерархических систем. При этом для наблюдения обычно доступно поведение систем лишь на верхних уровнях иерархии – землетрясения, результаты выборов и т.д.

Нелинейные иерархические системы не сводятся к простой сумме составляющих их частей; в них большую роль играет взаимодействие этих частей, приводящее к изменению свойств всей системы. Предполагается, что процессы, происходящие в системе, устроены подобным образом на различных ее иерар-

хических уровнях (масштабная инвариантность, самоподобие, или автомодельность).

На втором пути, в отличие от первого, ищется целостное, без особых деталей, описание системы на языке обобщенных свойств нелинейных динамических систем [1-6]. При таком подходе прогноз поведения основных особенностей поведения системы становится сравнительно долгосрочным, но без детальности характера развития событий. Однако, если описание составлено корректно, содержит набор необходимых прогностических свойств (правильно выбранное фазовое пространство) и позволяет устойчиво давать прогноз с приемлемой адекватностью (хорошим качеством), то его можно пытаться уточнять, переходя от общего к частному, т.е. рассматривать процесс на более низких иерархических уровнях в том месте и времени, на которые распространяется текущая тревога.

Именно второй путь позволил создать алгоритмы, дающие прогноз сильных землетрясений [7,8], прогноз результатов сенатских и президентских выборов в США [9,10], экономических рецессий и безработицы в США и Европе, а также прогноз экономических кризисов в России [11-13].

Изучение свойств временных рядов преступности г. Ярославля как одной из важных характеристик состояния социума привело к разработке методов прогноза скачков тяжких преступлений, излагаемых в данной статье.

Общность характера предкризисного поведения сложных динамических систем, в частности, социально-экономических систем, дает авторам основание воспринимать разработку таких подходов не как научную деятельность в различных областях, а как работу по созданию отдельных элементов общей системы мониторинга функционирования больших социумов. Такой мониторинг, при своей сравнительной простоте, позволил бы постоянно замерять “температуру” и “давление” в данной социальной системе и, в случае, когда эти параметры оказываются вне нормы, диагностировать заболевание, а, возможно, и предлагать методы лечения.

§1. Динамика преступности

Как отмечалось выше, развитие сильных событий в иерархических системах, как правило, подготавливается предварительной активизацией на предшествующих иерархических уровнях. Как подготовка сильных землетрясений, так и процесс трещинообразования, наблюдаемый в экспериментах по разрушению образцов под действием внешних напряжений, происходят согласно этой схеме. Соответственно, применительно к преступности, можно ожидать роста числа менее опасных преступлений перед вспышкой числа более тяжелых преступлений.

Как известно, распределения числа землетрясений по их энергии или трещин по их размеру имеют степенной характер. Такие распределения, будучи представлены в двойном логарифмическом масштабе, имеют линейный вид и в сейсмологии называются графиками повторяемости, характеризующимися их наклоном β . Попытаемся построить алгоритм прогноза скачков тяжелой пре-

ступности по поведению аналога графика повторяемости для недельной статистики преступлений и происшествий г. Ярославля за период 3.02.1993 – 25.06.2001.

В табл.1 приведен перечень основных типов преступлений (событий) с их кодировкой, соответствующее им число событий, и варианты используемых в статье способов разбиения разных видов преступлений на группы. Объединение разных видов преступлений в малое число групп обусловлено необходимостью генерализации данных по отдельным видам преступлений с целью получения устойчивых и статистически представительных результатов.

Рассмотрим некоторые специфические эффекты динамики преступности, проявления которых можно ожидать в рамках иерархических моделей развития динамических процессов.

Пронормируем каждый временной ряд числа событий данного вида преступлений на суммарное число случаев и на среднестатистическое отклонение. Тем самым будет достигнуто единообразие диапазона изменений числа событий по разным видам преступлений и происшествий.

Рассмотрим такие ряды в координатах: вид события – время и проинтер-

Таблица 1. Характеристика преступлений и происшествий в г. Ярославле за период с 3.02.1993 по 25.06.2001 (по классификации МВД).

Код	Типы преступления и происшествий	Число событий	Варианты разбиения на группы	
			1-й	2-й
01	Убийство	895	3	5
02	Нанесение тяжкого вреда здоровью	1 976	2	5
03	Телесные повреждения	24 457	1	3
04	Изнасилование	761	2	5
05	Разбойное нападение	1 058	2	5
06	Грабёж	10 044	2	4
07	Хулиганство	4 504	1	2
08	Кража	60 909	1	2
09	Хищения	86	1	3
10	Взяточничество	25	1	3
11	Фальшивомонетничество	1 525	1	2
13	Угон автотранспорта	2 593	1	1
14	Сопротивление сотрудникам милиции	534	1	3
15	Вымогательство	902	1	3
17	Мошенничество	2 621	1	3
18	Поджог	348	2	3
30	Прочее преступление	7 851	1	1
34	Дорожно-транспортное происшествие	4 571		1
35	Скоропостижная смерть	6 903		1
36	Пожар	2 345		1
37	Несчастный случай	2 264		1
38	Самоубийство	1 929		4
39	Обнаружение трупа	5 756	1	4
41	Безвестное исчезновение людей	5 078	1	5
60	Прочее происшествие	20 755		1

полируем полученную матрицу многочленами с помощью стандартных программных средств системы МатЛаб по этим координатам с целью размыть отдельные неповторяющиеся особенности и выделить более типичные.

При рассмотрении данных при осреднении от одной недели до месяца явных признаков процессов подготовки активизации более тяжких преступлений выявить не удалось. Наблюдаются лишь периоды синхронного роста числа преступлений и происшествий некоторых видов.

Ситуация меняется при больших осреднениях. На рис.1 приводятся интерполированные нормированные данные по 16 видам преступлений, полученные за непересекающиеся трехмесячные интервалы. Коды событий (табл.1) подразумевают уменьшение тяжести преступления с увеличением номера кода. Изменения в яркости от белого к черному соответствует изменению интенсивности преступлений от сильной к слабой. На рис.1 видна тенденция развития во времени неких цепочек активизации, ориентированных от менее тяжких преступлений к более тяжким. Отметим, что появление такого эффекта вряд ли может быть обусловлено исключительно характером использованного метода обработки данных.

Другой подход к выявлению активизации тяжких преступлений по пред-

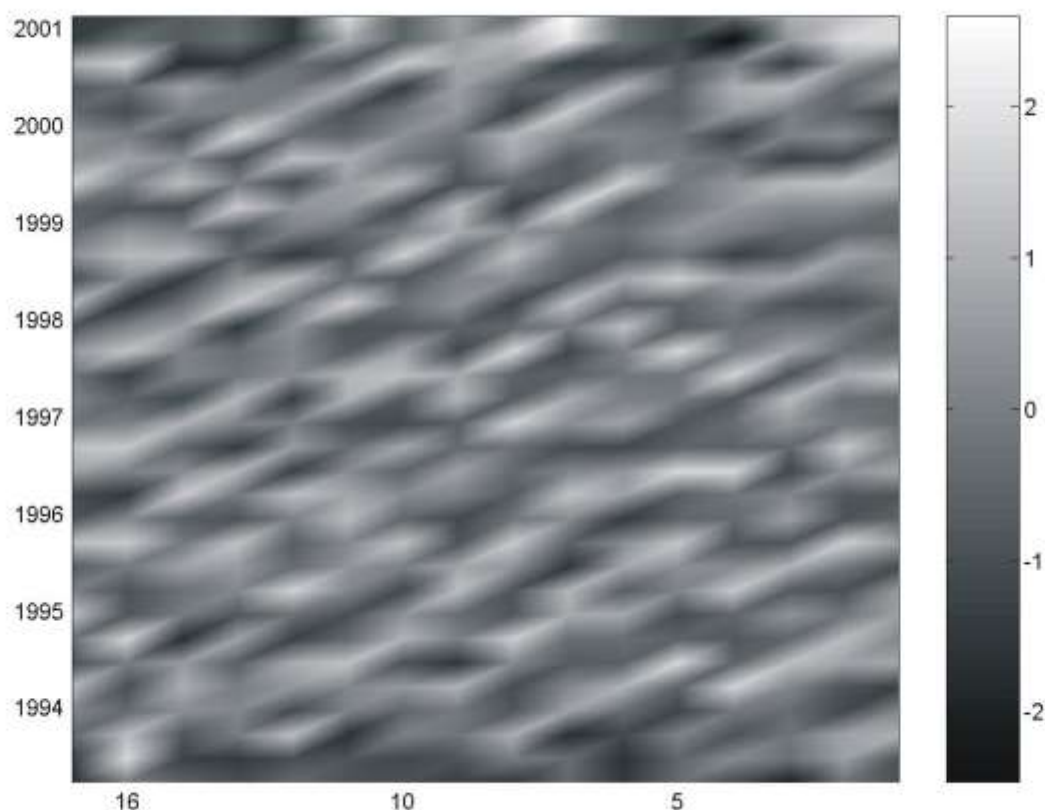


Рис. 1. Нормированные и интерполированные стандартным образом данные по режиму разных типов преступлений

По оси x – коды типов преступлений, по оси y – годы.

варяющим изменениям в режиме менее опасных видов преступлений основывается на анализе характера коррелированности (со сдвигом по времени) между разными видами преступлений. Для большей устойчивости и робастности будем рассматривать совокупности нескольких видов преступлений, объединяя преступления в группы. Нумерация групп ведется от групп менее тяжких преступлений к более тяжким, с тем, чтобы значения кодов более «сильных» преступлений были выше, как это принято при исследовании землетрясений и иных катастроф. Два варианта такого разбиения на группы приведены в табл.1.

Для каждого разбиения рассчитываются значения корреляции между рядом числа событий для группы самых тяжких преступлений и всеми группам разбиения, при этом ряд, соответствующий тяжким преступлениям брался с опережением от 1 до 15 недель.

Далее были определены величины изменения коррелированности по разным группам относительно их среднего значения и проведено сглаживание по трем соседним значениям времени запаздывания (например, запаздывание на две недели будет соответствовать среднему из исходных трех значений: для 1-й, 2-х и 3-х недель запаздывания). Полученные осредненные значения представлены на рис.2.

Из рис.2а и 2б можно видеть, что при такой высокой степени генерализации удастся выявить некоторые общие тенденции предвестникового режима преступлений. Во-первых, повышению числа тяжких преступлений с опережением 1-2 недели предшествует рост числа всех типов преступлений. Эту особенность режима естественно сопоставить с праздниками, некоторые из которых довольно продолжительны. Известно также, что праздники обычно сопровождаются ростом числа некоторых типов преступлений (в том числе и тяжких). Периоды синхронного роста числа преступлений разного типа выделялись также и при сглаживании режима преступности полиномами, аналогично ре-

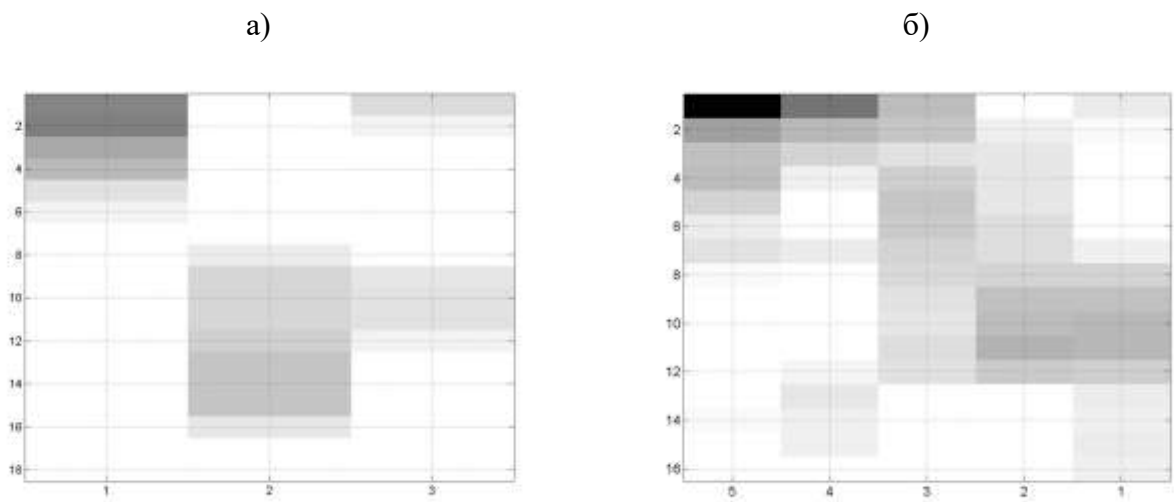


Рис. 2. Средние нормированные и сглаженные значения коррелированности для 1-го (а) и 2-го (б) вариантов разбиения преступлений и происшествий на группы

зультатам, приведенным на рис.1, но при временах сглаживания числа преступлений не более месяца.

Во-вторых, что более интересно, видно, что повышению числа наиболее тяжких преступлений с опережением месяц и более предшествуют периоды увеличения числа менее социально опасных видов преступлений. При этом волна активизации преступности как бы перетекает из области менее тяжких преступлений в правом нижнем углу рисунка в область более тяжких и более поздних преступлений в левом верхнем углу рисунка. Отметим, что этот результат рис.2 качественно аналогичен представленному ранее на рис.1. Близки по времени и интервалы времени предвестниковой активизации на рис.2 и длительности существования отдельных цепочек активизации на рис.1. Продолжительность этих предвестниковых проявлений варьируется от нескольких недель до полугода.

Типичной особенностью многих естественных и техногенных систем является наличие некоторой совокупности характерных периодов, которые выявляются в результате применения тех или иных методов спектрального или корреляционного анализа. Такие периоды могут накладываться на основную гладкую компоненту спектра, обычно близкую к белому спектру для равновесных систем или на розовый спектр с сильным доминированием низкочастотных компонент, что характерно для динамических систем. При этом считается, что стационарность набора характерных периодичностей отражает устойчивость

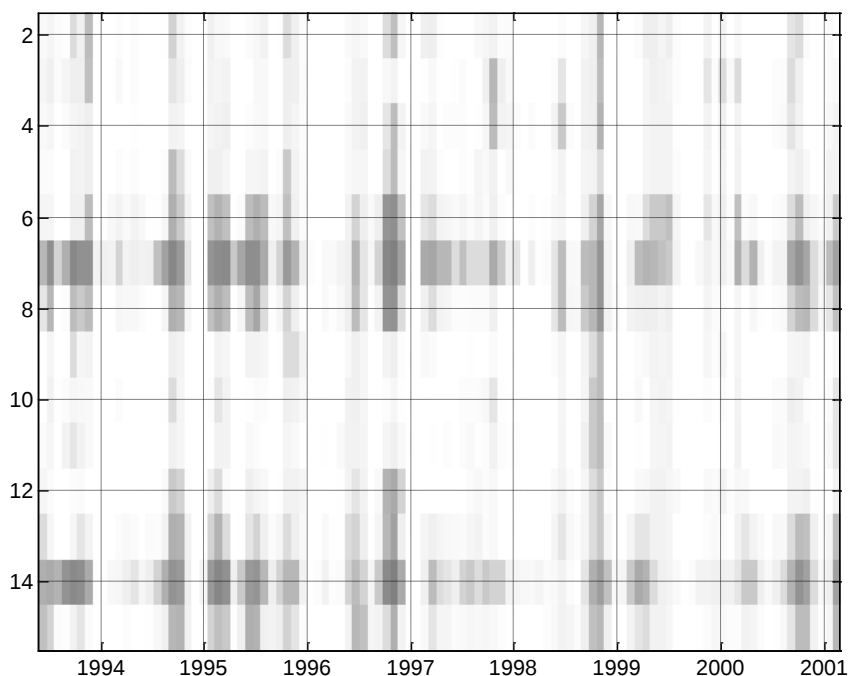


Рис. 3. Изменчивость периодической структуры режима недельного числа краж во времени

По оси x – время, по оси y – тоновое отображение величины автокорреляционной функции.

режима исследуемого процесса, т.е. в данном случае – временного режима числа преступлений. И, наоборот, исчезновение автокорреляционных максимумов или появление новых максимумов свидетельствует о существенных перестройках режима [14].

На рис.3. представлены данные по изменению автокорреляционной функции для временного ряда числа краж, самого многочисленного вида преступления. Значения функции автокорреляции, рассчитанные для скользящих временных интервалов длиной 3 месяца с временным сдвигом ряда ± 15 дней, даны градацией серого цвета, причем более темный цвет соответствует более высоким значениям. Значения автокорреляции при сдвиге на один день существенно больше остальных, возможно, из-за резкого отличия режима суббота-воскресений от остальных дней недели, поэтому на рис.3 приведены, для лучшей сопоставимости, результаты расчетов для сдвигов от 2 до 15 дней. По смыслу представляемой информации рис.3 в значительной степени аналогичен спектрально-временным СВАН диаграммам, широко используемым для исследования режима динамических систем разной природы [14,15].

На рис.3 видно, что на рассматриваемом интервале сдвигов явно и устойчиво доминируют недельная и 2-х недельная компоненты, но заметны также интервалы времени их исчезновения. Причем в эти интервалы времени обычно наблюдается быстрый рост или падение числа краж (рис.4). Например, регулярному исчезновению недельной периодичности числа краж в окрестности продолжительных новогодних праздников соответствует в большинстве случаев значительное уменьшение числа краж. Отметим, что потеря автокоррелированности ряда числа краж не является следствием сильной нестационарности

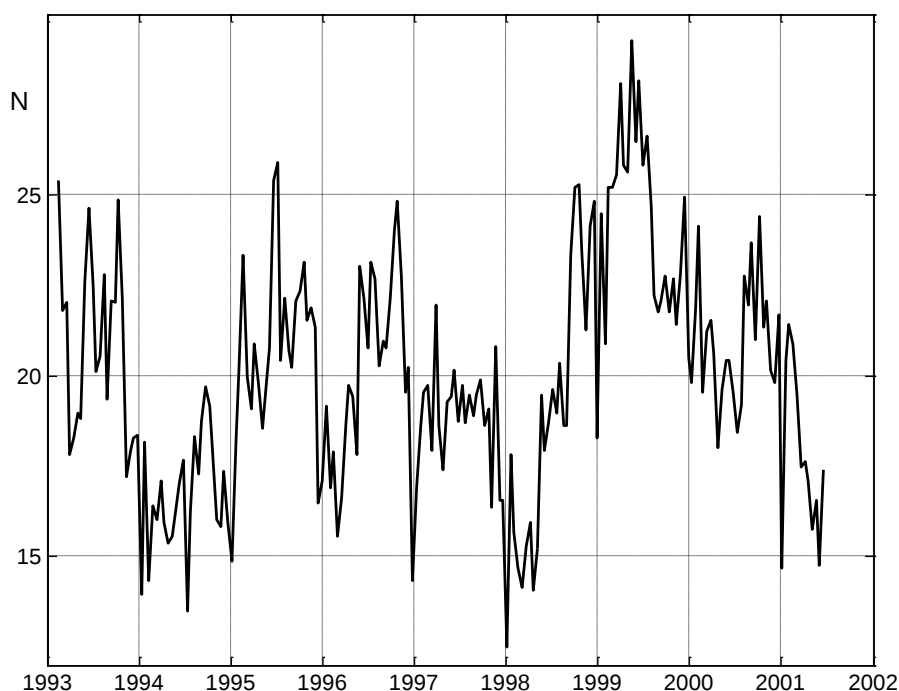


Рис. 4 Изменение во времени еженедельного числа краж

ряда в исследуемые интервалы времени, так как функция автокорреляции для всех интервалов времени рассчитывалась после удаления нелинейного (квадратичного) временного тренда.

Связывая оценки устойчивости режима преступлений с изменением социально-экономической обстановки, стоит обратить внимание на сильный сбой правильной периодической структуры в 1998 и начале 1999 года. Этому же интервалу времени соответствует быстрый рост числа краж (рис.4). Естественно связать эту особенность с резким ухудшением уровня жизни населения в период дефолта 1998 года.

Аналогичные изменения в окрестности дефолта получаются при анализе

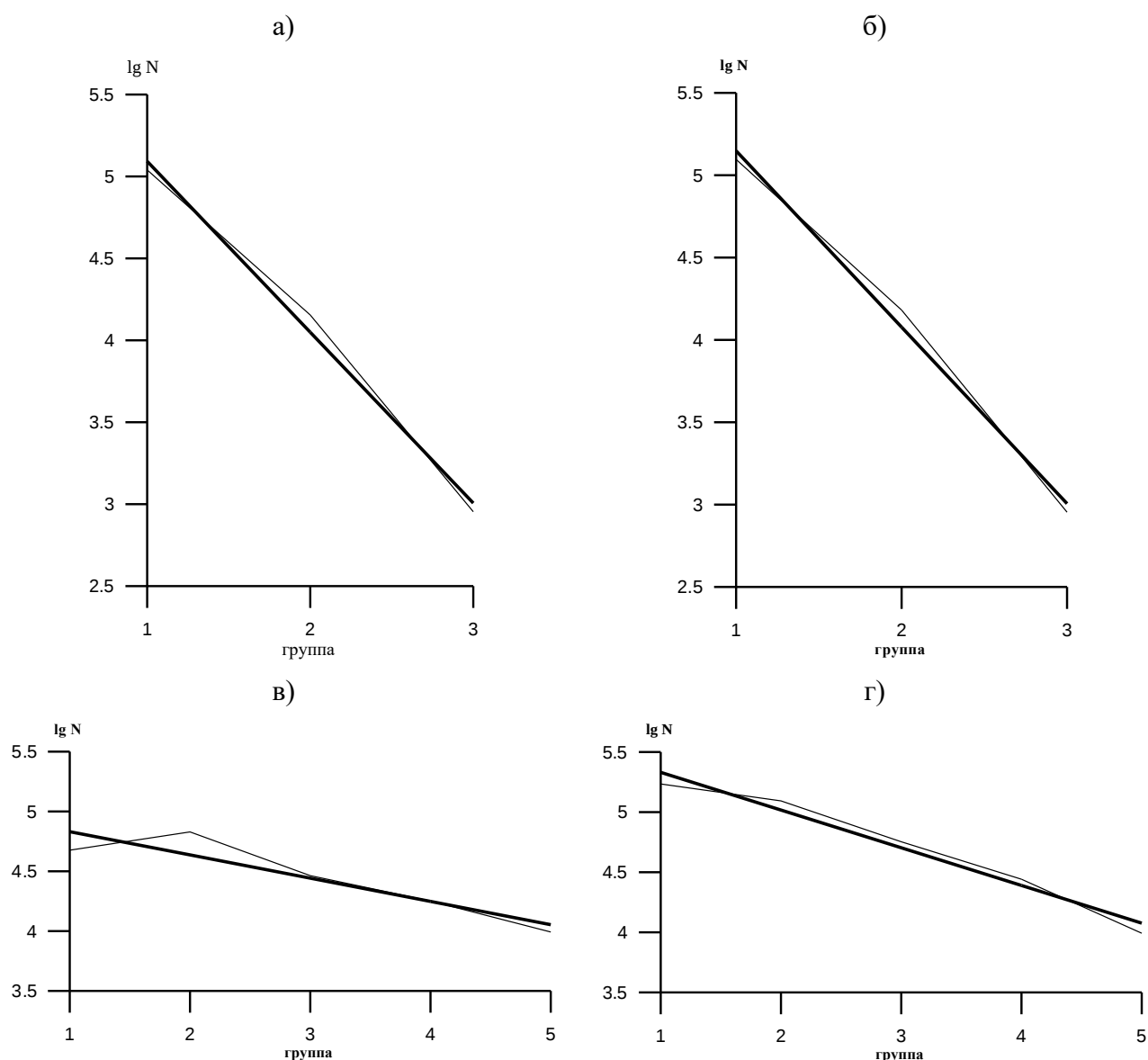


Рис. 5. Суммарные графики повторяемости для трех вариантов разбиения типов преступности в соответствии с табл. 1

Некумулятивные графики: а) – 1-й вариант разбиения, наклон $\beta = -1,0429$; в) – 2-й, $\beta = -0,1944$.

Кумулятивные графики: б) – 1-й варианта разбиения, наклон $\beta = -1,0710$; г) – 2-й, $\beta = -0,3136$.

других, достаточно представительных наборов данных, например группы 1, наименее тяжких преступлений 1-го варианта разбиения (табл.1).

§2. Прогноз скачков числа тяжких преступлений

Перенос методов прогноза физических явлений на социальные осложняется отсутствием метрических шкал измерения силы событий. В качестве аналога такого шкалирования типы преступлений и происшествий были разбиты на несколько групп. При этом, как это принято при анализе катастрофических явлений, группам более сильных событий, то есть более тяжких преступлений, присвоен более высокий ранг (табл.1).

Для каждого из разбиений при построении графика повторяемости возможны два способа представления – исходный и *кумулятивный*, когда в данную группу К входят также и события большей тяжести. Например, для 1-го разбиения первую кумулятивную группу (К1) составят события исходных (некумулятивных) групп Г1, Г2 и Г3, вторую (К2) – события групп Г2 и Г3, а третья (К3) будет совпадать с Г3. По аналогии с результатами исследования сейсмического режима можно предполагать, что применение кумулятивного способа представления обеспечит большую устойчивость результатов, возможно за счет некоторого ухудшения чувствительности.

Графики повторяемости за весь указанный период для некумулятивных и кумулятивных разбиений даны на рис.5.

Объекты прогноза. Объектами прогноза при этом будут моменты времени, на которые приходятся большие скачки числа событий группы, соответ-

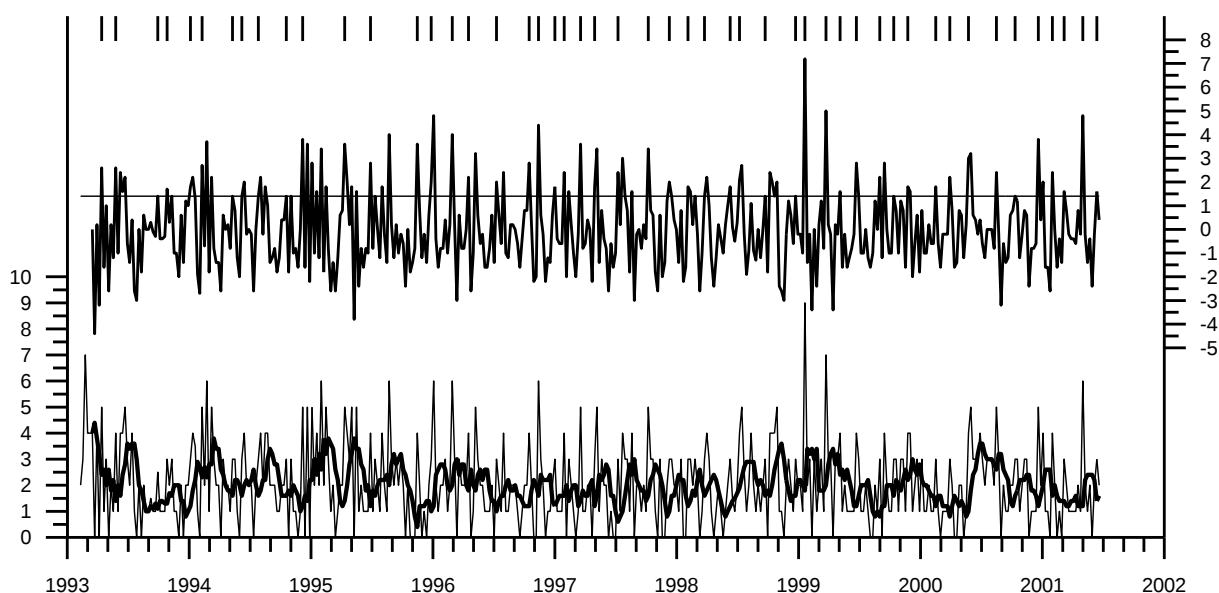


Рис. 6. Определение объектов прогноза – скачков тяжелой преступности – по ряду остатков (способ S) для 1-го варианта разбиения

Внизу (левая вертикальная ось) – исходный ряд числа событий Г3 (тонкая линия) и сглаженный 5-недельным окном (жирная линия); в середине (правая вертикальная ось) – разница между значениями исходного (Г3) и сглаженного рядов и порог отсечения $\sigma^* = 1,4$ (горизонтальная линия);верху – объекты прогноза.

ствующей тяжким преступлениям (ГЗ – для 1-го варианта разбиения и Г5 – для 2-го). Определение скачков может быть проведено следующим *способом S*.

Для каждой точки исходного ряда тяжких преступлений, начиная с шестой, вычисляется среднее за пять предыдущих измерений (недель) – то есть проводится сглаживание ряда 5-недельным окном. Составляется ряд остатков, состоящий из разницы между числом событий за данную неделю и полученным средним значением. *Большими* называются значения ряда остатков, не меньшие заданного *порога отсечения* σ^* . Если большие значения отстоят друг от друга менее чем на два шага измерения, из них оставляется лишь первое. Моменты времени, соответствующие оставшимся большим значениям и будут объектами прогноза (рис.6).

Построение прогнозной функции F1. Для каждой точки наблюдений, начиная с пятой, строится график повторяемости событий, произошедших за данную и четыре предыдущих недели. Если используется 1-й вариант разбиения, то графики повторяемости строятся для всех его трех групп Г1-Г3, а в случае 2-го – лишь для старших групп ГЗ-Г5, так как они более стационарны во времени.

Определим наклон графика β как модуль тангенса угла наклона аппроксимирующей прямой по методу наименьших квадратов. Как и при определении объектов, сгладив 5-недельным окном ряд значений β , получим ряд остатков F1, состоящий из разницы между текущим и сглаженным значением β . Если ввести теперь порог отсечения β^* , то достижение или превышение этого порога значениями F1 и будет считаться прогнозным сигналом. На рис.7, соответствующий

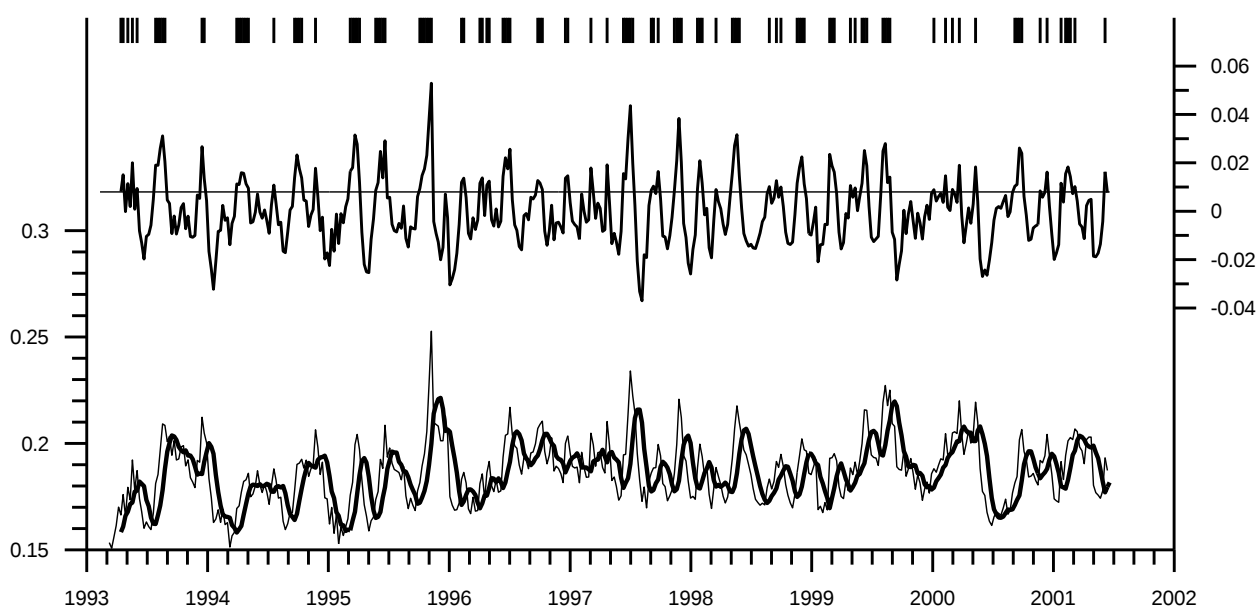


Рис.7. Построение F1 – прогнозной функции по остаткам – для 1-го варианта разбиения

Внизу (левая ось) – ряд текущих значений наклона некумулятивного графика повторяемости β для Г1-Г3 за 5 недель (тонкая линия) и ряд β , сглаженный 5-недельным окном (жирная линия); в середине (правая ось) – ряд остатков F1, состоящий из разницы между текущим и сглаженным значением β , и порог отсечения $\beta^*=0,008$; сверху – моменты появления прогнозного сигнала.

ющем данным рис.6, приведен пример построения прогнозной функции $F1$ для 1-го варианта разбиения.

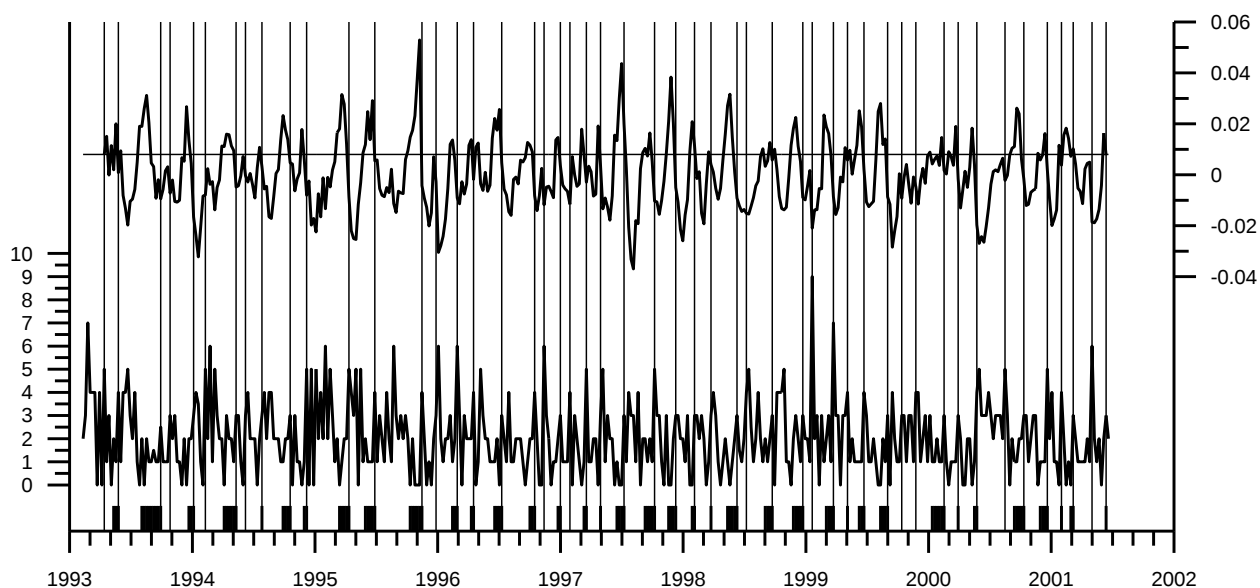


Рис.8. Прогноз при выборе объектов по способу S с применением прогнозной функции $F1$ для 1-го варианта разбиения ($\sigma^*=1,4$; $\beta^*=0,008$; $\Delta=6$)

Внизу – исходный ряд ГЗ (левая вертикальная ось);верху – ряд остатков $F1$ по некумулятивному графику повторяемости для Г1-ГЗ (правая вертикальная ось) и порог отсечения β^* (горизонтальная линия); тонкие вертикальные линии – объекты прогноза; черные прямоугольники – интервалы тревоги.

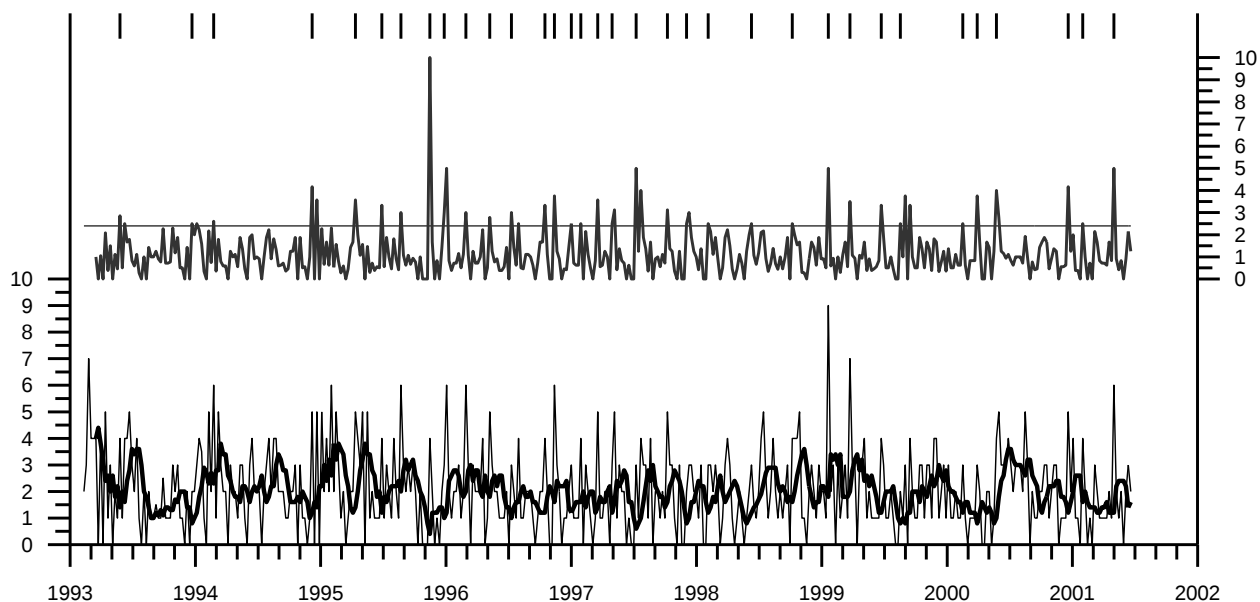


Рис.9. Определение объектов прогноза – скачков тяжелой преступности – по ряду отношений (способ D) для 1-го варианта разбиения

Внизу (левая вертикальная ось)– исходный ряд числа событий ГЗ (тонкая линия) и сглаженный 5-недельным окном (жирная линия); в середине (правая вертикальная ось) – отношение значений исходного ряда (ГЗ) к значениям сглаженного и порог отсечения $\sigma^* = 2,4$ (горизонтальная линия);верху – объекты прогноза.

Объявления тревоги. С момента первого появления прогнозного сигнала происходит *объявление тревоги* и она может длиться Δ недель, если за это время не возникает объект прогноза. В последнем случае *тревога заканчивается*. Если за время тревоги объектов нет, а обнаружен очередной прогнозный сигнал, то от этого момента *тревога продлевается* еще на Δ недель. Если объект и прогнозный сигнал возникают одновременно или сигнал отстоит от объекта на 1-2 недели, то тревога не объявляется.

Все тревоги, содержащие объекты прогноза, считаются *успешными*, а не содержащие – *ложными*.

Все объекты, попавшие в тревоги считаются *предсказанными*, а не попавшие считаются *пропусками цели*.

На рис.8 в графическом виде представлены результаты работы алгоритма прогноза с использованием выбора объектов по ряду остатков и с применением прогнозной функции $F1$ для 1-го варианта разбиения при $\sigma^* = 1,4$; $\beta^* = 0,008$ и $\Delta = 6$, соответствующие рис.6 и 7. В данном случае предсказано 37 из 50 объектов (74%), при этом суммарно тревога занимает 125 из 428 недель, что составляет 29,2% времени наблюдения при отсутствии ложных тревог (а длительность 2-х ложных тревог составляет 8,3% от суммарного времени всех тревог).

При определении объектов *способом D* вместо ряда остатков, используемого при способе S, рассматривается ряд отношений значений исходного ряда к сглаженному.

Для 1-го варианта группирования определение объектов таким способом

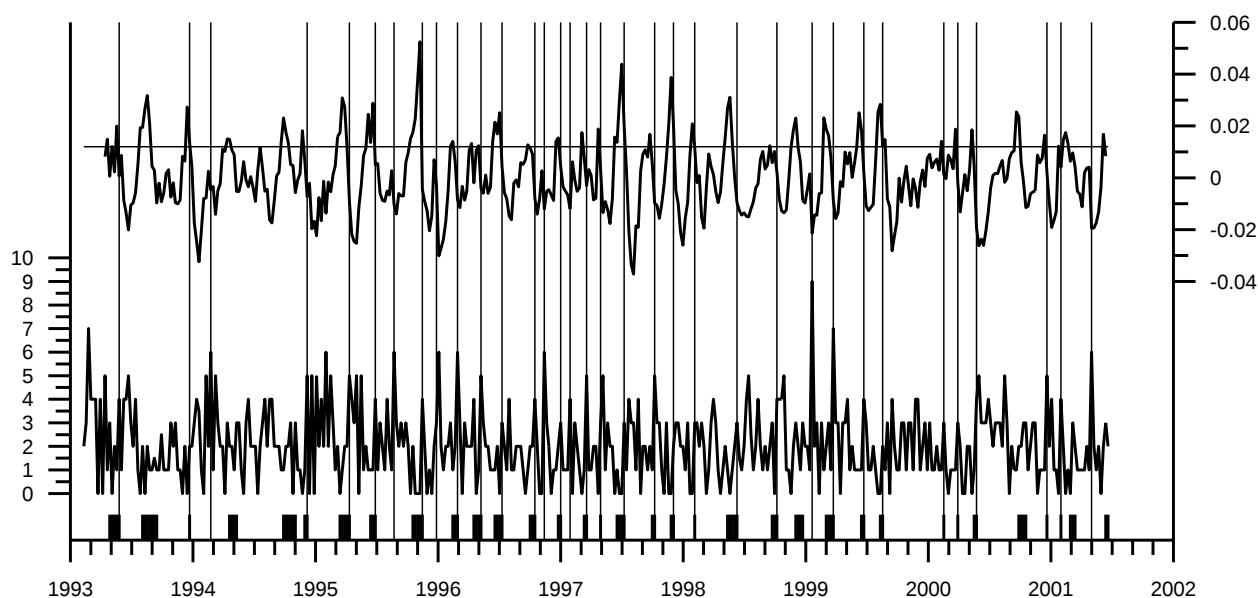


Рис.10. Прогноз объектов, полученных способом D (рис.9) для 1-го варианта разбиения с применением прогнозной функции $F1$ по кумулятивному графику повторяемости ($\sigma^*=2,4$; $\beta^*=0,012$; $\Delta=3$)

Внизу – исходный ряд ГЗ (левая вертикальная ось);верху – ряд остатков $F1$ по графику повторяемости для К1-К3 (правая вертикальная ось) и порог отсечения β^* (горизонтальная линия); тонкие вертикальные линии – объекты прогноза; черные прямоугольники – интервалы тревоги.

показано на рис.9, а на рис.10 приведены результаты работы алгоритма прогноза для этих объектов с применением прогнозной функции F1 по кумулятивному представлению графика повторяемости при $\sigma^*=2,4$; $\beta^*=0,012$ и $\Delta=3$. Здесь прогнозируется 27 из 34 объектов (79%), при этом тревога составляет 23,6% времени наблюдения, а длительность 6 ложных тревог составляет 29,7% от суммарного времени тревоги.

Построение прогнозной функции F2. При построении другой прогнозной функции F2, для каждой точки наблюдений, начиная с пятой, строится график повторяемости событий, произошедших за данную и четыре предыдущих недели и снова определяется наклон этого графика, но для всех групп событий выбранного разбиения. Условием, при котором F2 включает прогнозный сигнал, является наличие таких пяти последовательных моментов времени, отстоящих друг от друга не более чем на три шага измерения, что значения ряда F2 последовательно возрастают; причём значения предиктора в моменты времени, находящиеся между двумя соседними точками из выбранных пяти, должны быть не больше значений предиктора в этих двух соседних точках.

Выбор объектов способом D при $\sigma^* = 1,6$, построение прогнозной функции F2 по кумулятивному графику повторяемости ($\Delta=3$) и результаты прогноза представлены на рис.11–13 для 2-го варианта разбиения на группы. В этом случае прогнозируется 10 из 12 объектов (83%) при общем времени тревоги 22,4%, а длительность 14 ложных тревог составляет 58,8% от суммарного времени всех тревог. Данный результат нельзя считать хорошим, поскольку число ложных тревог слишком велико.

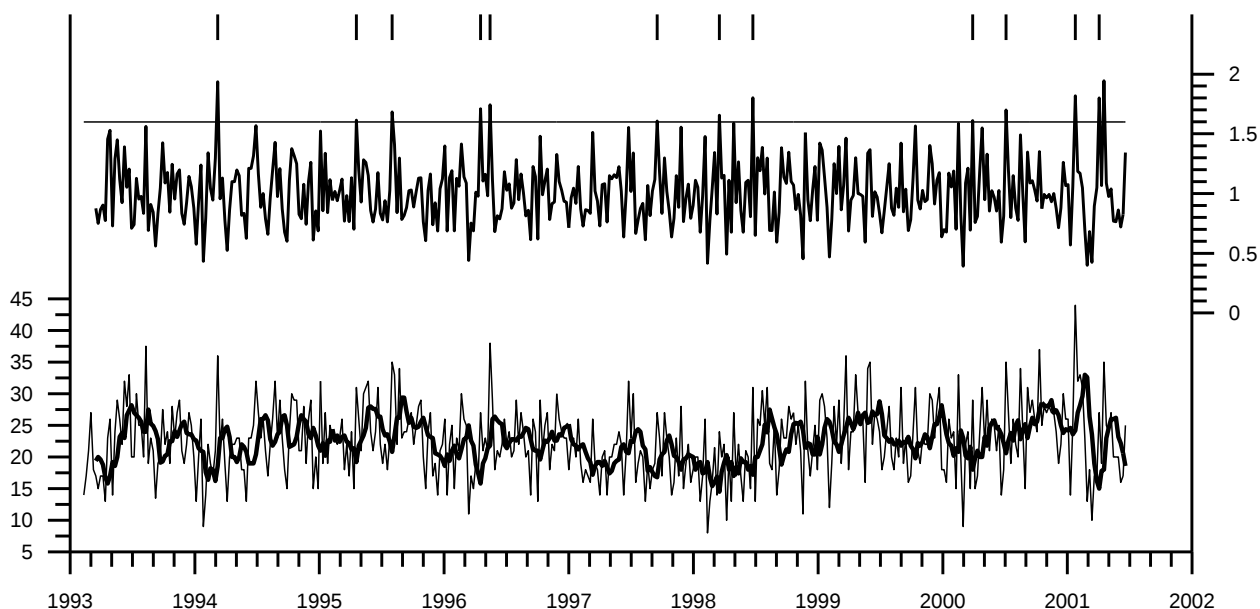


Рис.11 Определение способом D объектов прогноза для 2-го варианта разбиения

Внизу (левая вертикальная ось) – исходный ряд числа событий Г5 (тонкая линия) и сглаженный 5-недельным окном (жирная линия); в середине (правая вертикальная ось) – отношение значений исходного ряда (Г5) к значениям сглаженного и порог отсечения $\sigma^* = 1.6$ (горизонтальная линия);верху – объекты прогноза.

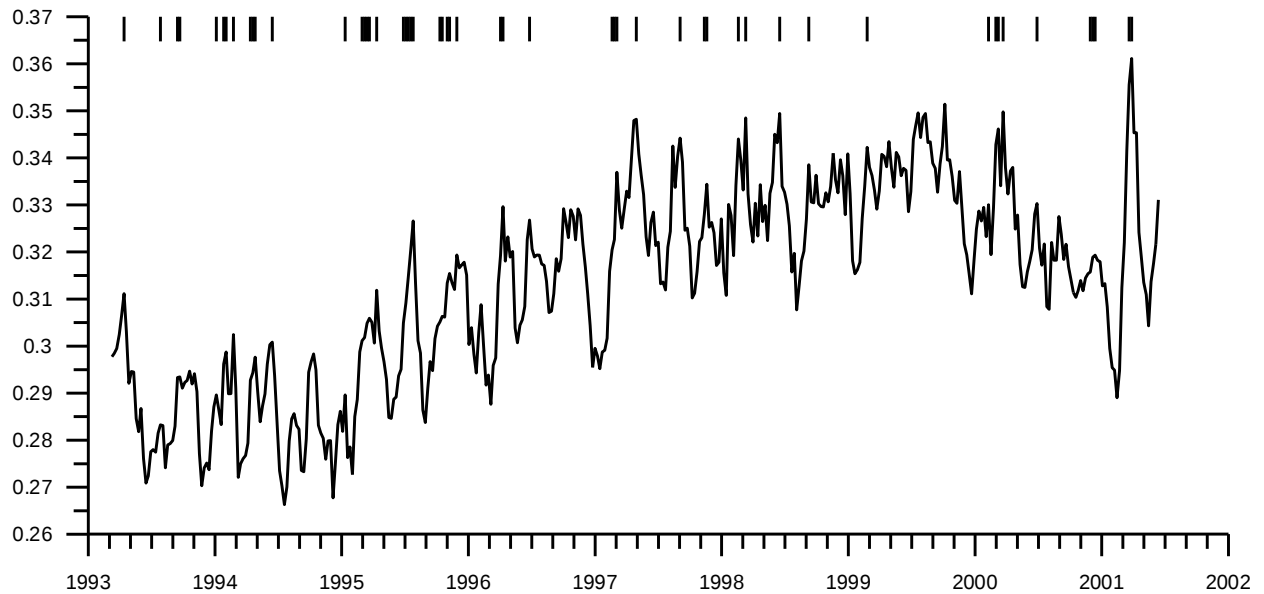


Рис.12. Построение прогнозной функции F2 для 2-го варианта разбиения по кумулятивному графику повторяемости

Внизу – ряд F2 текущих значений наклона кумулятивного графика для K1-K5 за 5 недель; вверху – моменты появления прогнозного сигнала.

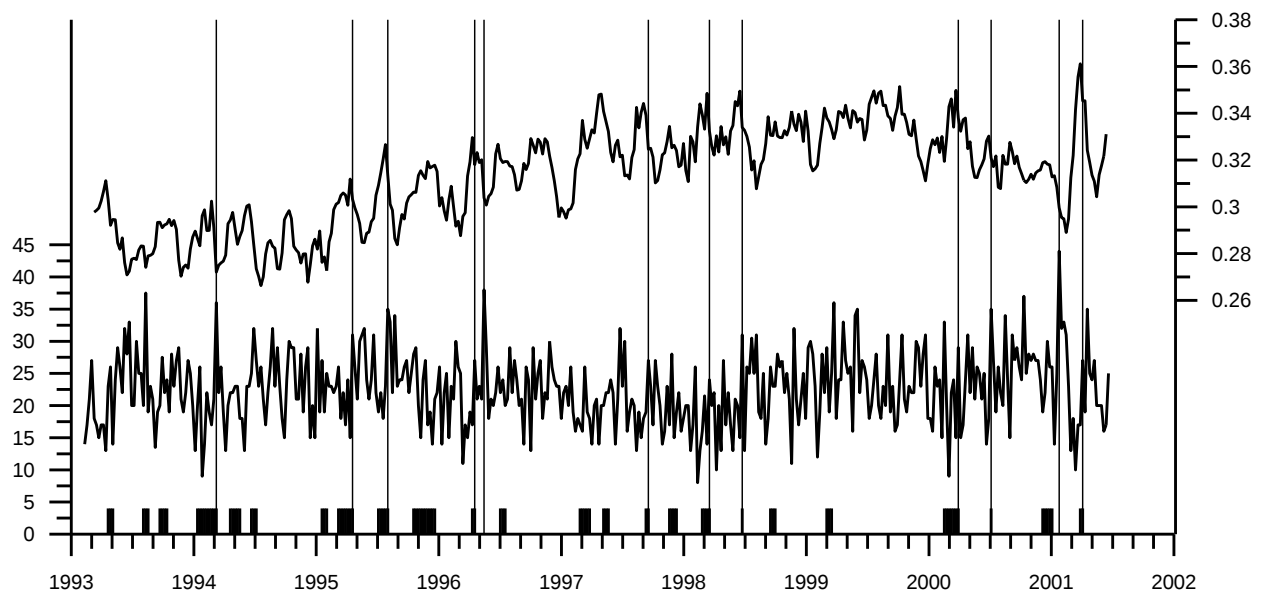


Рис.13. Прогноз объектов, полученных способом D (рис.11) для 2-го варианта разбиения с применением прогнозной функции F2 по кумулятивному графику повторяемости ($\sigma^*=1.6$, $\Delta=3$)

Внизу – исходный ряд Г5 (левая вертикальная ось); вверху – ряд значений наклона кумулятивного графика повторяемости для K1-K5 (правая вертикальная ось); тонкие вертикальные линии – объекты прогноза; черные прямоугольники – интервалы тревоги.

Численные эксперименты по проверке алгоритмов прогноза. Итак, были предложены следующие возможные параметры алгоритма прогноза скачков тяжкой преступности:

- два варианта разбиения типов преступности и происшествий на группы;

- два варианта выбора объектов – по остаткам и по отношению (S и D), с порогом отсечения σ^* ;
- два варианта представления графика повторяемости – некумулятивный и кумулятивный (nK и K);
- два вида прогнозной функции – $F1$ с порогом отсечения β^* и $F2$;
- длительность тревоги Δ .

Обозначим число всех объектов через N , а предсказанных, т.е. попавших в интервал тревоги, – через N^+ . Тогда число пропусков цели, т.е. объектов вне интервала тревоги составит $N^- = N - N^+$, а $\eta = N^-/N$. Если за все время T суммарное время тревоги составило T_a , то $\tau = T_a/T$.

Интегральное качество прогноза удобно оценивать по η - τ диаграмме [16,8]. При $\eta = 1$, $\tau = 0$ имеем предельный случай, когда пропускаются все события при времени тревоги, равном нулю («стратегия оптимиста»), а при $\eta = 0$, $\tau = 1$ предсказываются все события при постоянной тревоге («стратегия пессимиста»). Интегральное качество прогноза при изменениях свободных параметров алгоритма можно оценить величиной $\varepsilon = \min(\eta + \tau)$ или, при разумных изменениях параметров, его средним. Условие $\eta + \tau < 1$ ($\varepsilon < 1$) соответствует нетривиальному прогнозу и чем меньше ε , тем лучше прогноз. Для прогноза землетрясений значения ε находятся в пределах 0.32-0.5. Приведенный выше (рис.13) пример прогноза скачков преступности дает $\varepsilon = 0.39$.

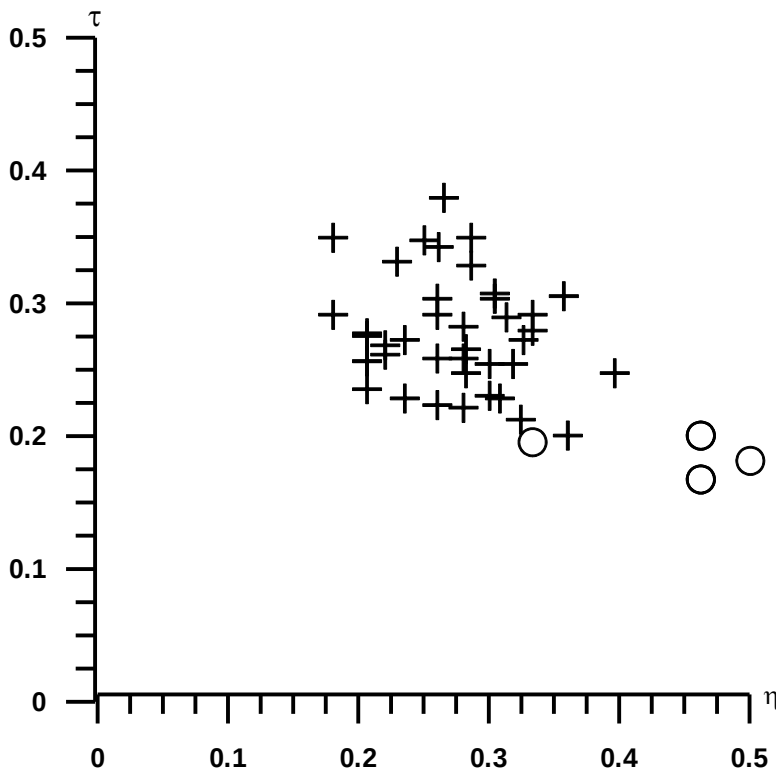


Рис. 14. η - τ диаграмма результатов прогноза, приведенных в табл. 2

“+” – результаты из табл.2а, “о” – из табл.2в.

Добавим еще два показателя качества алгоритма – число ложных тревог и их длительность Tf / T_a относительно суммарного времени тревоги. Вместе с η и τ из них можно было бы создать 4-х мерное пространство метрик и в нем ввести аналог ε . Но в данной работе для различных вариантов прогноза мы будем лишь приводить значения ε , числа ложных тревог и Tf / T_a .

Результаты прогноза при различных параметрах алгоритма для случаев, когда число ложных тревог составляет менее 30% от числа всех объектов прогноза представлены на суммарной η - τ диаграмме рис.14 и приведены в табл.2а и 2в.

Кроме того был проведен эксперимент по исключению из рассмотрения моментов времени, приходящихся на праздничные дни каждого года: с 30 декабря по 20 января, с 23 февраля по конец февраля, с 7 по 14 марта, с 30 апреля по 15 мая, с 11 по 18 июня, с 6 по 13 ноября и с 11 по 18 декабря. Результаты этого эксперимента приведены в табл.2б и 2г.

Таблица 2. Суммарная сводка результатов прогноза при различных параметрах
а) Результаты прогноза при применении функции F1

№/пп	S/D	nK/K	σ^*	β^*	Δ	$N^+:N$	η	τ	ε	N_f	Tf/T_a
1-й вариант разбиения на группы											
1	S	nK	1,4	0,008	3	36:50	,280	,283	,563	2	0,083
2	S	nK	1,4	0,008	6	37:50	,260	,292	,552	0	0
3	S	nK	1,4	0,010	3	35:50	,300	,255	,555	1	0,064
4	S	nK	1,4	0,010	6	36:50	,280	,259	,539	0	0
5	S	nK	1,6	0,008	3	32:46	,304	,304	,608	6	0,277
6	S	nK	1,6	0,008	6	34:46	,261	,343	,604	2	0,129
7	S	nK	1,6	0,010	3	31:46	,326	,273	,599	5	0,274
8	S	nK	1,6	0,010	6	32:46	,304	,308	,612	2	0,220
9	D	nK	2,4	0,011	3	27:34	,206	,257	,463	6	0,327
10	D	nK	2,4	0,012	3	26:34	,235	,229	,464	7	0,337
11	D	nK	2,4	0,012	6	26:34	,235	,273	,508	5	0,333
12	D	K	2,0	0,010	3	30:44	,318	,255	,573	5	0,284
13	D	K	2,2	0,010	3	28:39	,282	,266	,548	6	0,316
14	D	K	2,2	0,011	3	28:39	,282	,248	,530	6	0,321
15	D	K	2,2	0,012	3	27:39	,308	,229	,537	6	0,306
16	D	K	2,4	0,012	3	27:34	,206	,236	,442	6	0,297
17	D	K	2,4	0,012	6	27:34	,206	,278	,484	5	0,327
18	D	K	2,4	0,010	3	27:34	,206	,276	,482	7	0,331
19	D	K	2,4	0,011	3	27:34	,206	,257	,463	7	0,336
20	D	K	2,4	0,013	3	23:34	,324	,213	,537	8	0,396
2-ой вариант разбиения на группы											
21	D	nK	1,1	0,000	3	41:50	,180	,350	,530	6	0,200
22	D	nK	1,1	0,005	3	37:50	,260	,304	,564	6	0,246
23	D	nK	1,1	0,010	3	35:50	,300	,231	,531	5	0,232
24	D	nK	1,1	0,010	6	37:50	,260	,259	,519	2	0,126
25	D	nK	1,1	0,015	3	32:50	,360	,201	,561	5	0,267
26	D	nK	1,2	0,005	3	36:48	,250	,348	,598	5	0,195
27	D	nK	1,2	0,010	3	32:48	,333	,292	,625	5	0,208
28	D	nK	1,2	0,015	3	29:48	,396	,248	,644	5	0,255
29	D	nK	1,3	0,010	3	30:42	,286	,350	,636	7	0,267
30	D	nK	1,3	0,015	3	27:42	,357	,306	,663	8	0,305
31	D	nK	1,4	0,010	3	25:34	,265	,380	,645	11	0,380
32	D	K	1,1	0,006	3	39:50	,220	,269	,489	5	0,209
33	D	K	1,1	0,010	3	37:50	,260	,224	,484	6	0,250
34	D	K	1,1	0,010	6	39:50	,220	,262	,482	3	0,196
35	D	K	1,1	0,005	3	41:50	,180	,292	,472	6	0,240
36	D	K	1,2	0,006	3	37:48	,229	,332	,561	7	0,282
37	D	K	1,2	0,010	3	33:48	,313	,290	,603	7	0,306
38	D	K	1,1	0,011	3	36:50	,280	,222	,502	6	0,252
39	D	K	1,2	0,011	3	32:48	,333	,280	,613	7	0,292
40	D	K	1,3	0,011	3	30:42	,286	,329	,615	7	0,255

б) Результаты прогноза при применении функции F1 для рядов без праздников

№/пп	№ в 2а	S/D	σ^*	nK/K	β^*	Δ	$N^+:N$	η	τ	ε	N_f	Tf/T_a
1-й вариант разбиения на группы												
1	1	<u>S</u>	1,4	nK	0,008	3	30:41	,268	,293	,561	6	0,172
2	2	S	1,4	nK	0,008	6	31:41	,244	,305	,549	4	0,087
3	4	S	1,4	nK	0,010	6	30:41	,268	,272	,540	0	0
4	9	D	2,4	nK	0,011	3	24:28	,143	,263	,406	4	0,337
5	11	D	2,4	nK	0,012	6	24:28	,143	,272	,415	4	0,402
6	17	D	2,4	K	0,012	6	24:28	,143	,278	,421	4	0,394
2-ой вариант разбиения на группы												
7	23	<u>D</u>	1,1	nK	0,010	3	28:41	,317	,237	,554	4	0,263
8	24	D	1,1	nK	0,010	6	29:41	,293	,263	,556	4	0,101
9	32	D	1,1	K	0,006	3	31:41	,244	,263	,507	4	0,157
10	34	D	1,1	K	0,010	6	31:41	,244	,260	,504	3	0,114

в) Результаты прогноза при применении функции F2

№/пп	S/D	σ^*	nK/K	Δ	$N^+:N$	η	τ	ε	N_f	Tf/T_a
1-й вариант разбиения на группы										
1	D	2,6	nK	3	14:26	,462	,168	,630	7	0,431
2	D	2,6	nK	5	14:26	,462	,201	,662	7	0,523
3	D	2,6	K	3	14:26	,462	,168	,630	7	0,458
4	D	2,6	K	5	14:26	,462	,201	,662	7	0,535
2-ой вариант разбиения на группы										
5	D	1,4	K	3	17:34	,500	,182	,682	7	0,410
6	D	1,5	K	3	16:24	,333	,196	,530	8	0,429

г) Результаты прогноза при применении функции F2 для рядов без праздников

№/пп	№ в 2в	S/D	σ^*	nK/K	Δ	$N^+:N$	η	τ	ε	N_f	Tf/T_a
1-й вариант разбиения на группы											
1	1	D	2,6	nK	3	13:20	,350	,180	,530	5	0,377
2	2	D	2,6	nK	5	13:20	,350	,216	,566	4	0,479
3	3	D	2,6	K	3	13:20	,350	,178	,528	6	0,400
4	4	D	2,6	K	5	13:20	,350	,216	,566	4	0,507

Из табл.2 видно, что, в среднем, лучше всего работает функция F1 для 1-го варианта разбиения, метода определения объектов по остаткам (S) и некумулятивного графика повторяемости. Для этой группы вариантов существует интервал параметров алгоритма, в которых он дает достаточно хорошие и устойчивые результаты – $\sigma^*=1,4-1,6$; $\beta^*=0,008-0,010$; $\Delta=3-6$.

Результаты проверки алгоритмов для разных значений используемых параметров позволили выявить между ними ряд взаимосвязей, учет которых может представлять интерес при применении данных алгоритмов прогноза. Так, при уменьшении порога выбора объектов прогноза σ^* ниже 1,4 число объектов прогноза увеличивается, и среднее расстояние между ними уменьшается. Так как при определении объектов прогноза, когда объекты отстоят друг от друга менее чем на два шага измерения, более позднее значение за объект прогноза

нами не принимается, то имеем, что число не группирующихся объектов прогноза резко уменьшается. Отсюда видно, что для задач прогноза с малыми превышениями величин объектов прогноза над фоном предложенный алгоритм должен быть изменен.

При уменьшении порога объявления тревоги β^* ниже 0,008 общее время тревоги начинает превышать 30% от всего времени наблюдения, для практических применений столь большое время тревоги представляется неудобным, результаты этих расчетов в таблицу не вошли.

В случае выхода за верхние границы соответствующих оптимальных интервалов для σ^* и β^* сокращается число объектов прогноза, но при этом увеличивается число ложных тревог, превышая введенный нами порог 30% от числа всех объектов прогноза.

Исключение праздничных дней, в среднем, слегка улучшает качество прогноза.

Полученные результаты прогноза формально достаточно хороши, но не вполне понятен механизм реализации прогнозных изменений режима преступности. Из поведения используемых прогнозных функций видно, что предвестником является увеличение угла наклона β графика повторяемости, что, казалось бы, противоречит нашим исходным посылкам.

Возьмем все объекты прогноза из примера на рис.6 и углы наклона β соответствующего сглаженного 5-недельным окном графика повторяемости, использующегося при построении функции F1 (рис.7). Рассмотрим средние значения β как в точке появления объекта, так и в 5 предшествующих точках (рис.15). Такой способ рассмотрения называется методом наложения эпох, в нашем случае таким путем получаем средний сценарий подготовки скачка числа тяжких преступлений.

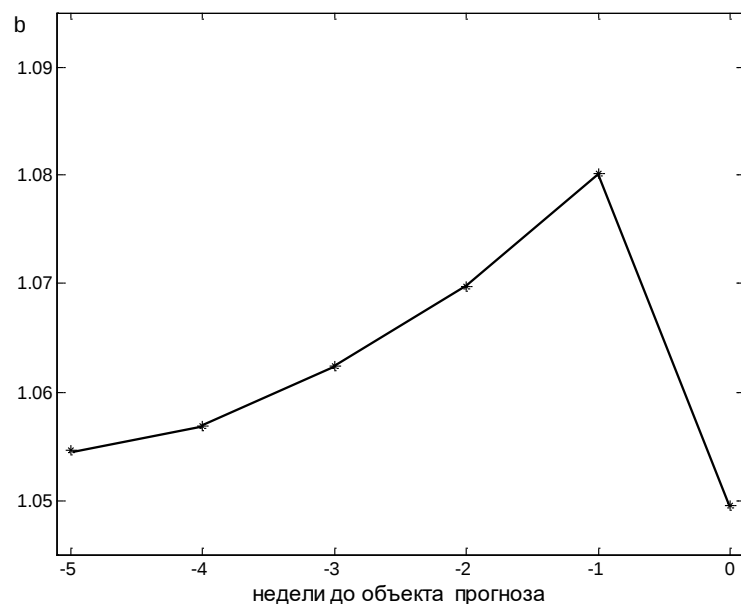


Рис.15. Изменения среднего угла наклона графика повторяемости β за 5 недель, предвещающих скачки тяжкой преступности для 1-ого варианта разбиения при выборе объектов по способу S (рис.6 и 7)

На графике хорошо виден использовавшийся выше прогнозный критерий увеличения угла наклона графика повторяемости. Падение значения β вниз в

последней точке графика связано с самим объектом прогноза, то есть скачком числа тяжких преступлений, что должно отвечать сильному уменьшению наклона графика повторяемости.

Однако наклон графика повторяемости является интегральной характеристикой изменения числа событий разных групп – как сильных, так и слабых преступлений. Поэтому рассмотрим теперь характер предваряющих изменений в каждой из этих групп от 1-й (происшествия) до 3-й (убийства), по отдельности. Для удобства сопоставления характера изменения числа преступлений разных типов, приводимые изменения отнормированы на среднее число преступлений соответствующих типов.

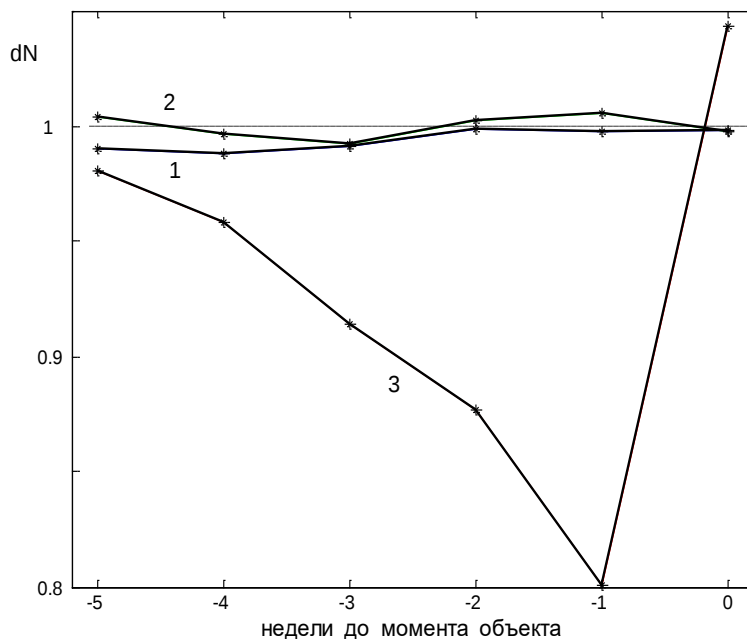


Рис.16. Изменения нормированного среднего числа событий 1, 2, и 3-й групп, предваряющих скачки тяжелой преступности для 1-го варианта разбиения при выборе объектов по способу S (рис.6 и 7)

Для групп 1 и 2 существенных закономерных изменений предваряющих изменений числа событий, кроме тенденции слабого роста числа событий разных типов за 1-2 недели до большого всплеска активности (предположительно упоминавшейся ранее «эффект больших праздников»), не наблюдается. Существенно иначе ведет себя число событий 3-й группы – убийств; оно прогрессивно уменьшается вплоть до момента появления объекта прогноза. Именно этот эффект и приводит

к увеличению наклона графика повторяемости как прогнозного признака. Полученный сценарий предвестниковых изменений режима тяжких преступлений допускает несколько возможных интерпретаций, причем возможен одновременный вклад всех из них. Выявленный механизм реализации прогноза может рассматриваться как аналог явления сейсмического «затишья», наблюдаемого перед некоторыми сильными землетрясениями. Отчасти такой характер предвестниковых изменений может быть следствием выбора объектов прогноза, как скачков числа событий относительно некоторого предыдущего уровня. Возможно также, что число групп в разбиениях недостаточно велико для выявления отчетливой картины «прорастания» активности от нижних уровней

иерархической системы преступности к верхним, как это можно видеть, при большой степени генерализации, на рис.1 и 2.

ЗАКЛЮЧЕНИЕ

Выше было продемонстрировано, что режим изменения числа преступлений во многом аналогичен поведению динамических систем с эффектами предвестниковой активизации и затишья перед сильными скачками числа тяжких преступлений. При этом явление активизации довольно продолжительно – от нескольких месяцев до полугода, и проявляется в «проращении» менее опасных преступлений и происшествий в более тяжкие.

Эффект затишья более локализован во времени – продолжается не более полутора месяцев и связан с изменением числа тяжких преступлений. По аналогии с моделью развития экстремально сильных событий [17], реализация тяжких преступлений может трактоваться как совокупность эффектов создания неравновесного состояния системы и некоего спускового механизма. Соответственно, одним из объяснений предвестникового затишья в режиме тяжких преступлений может быть «эффект задержки», когда предпосылки для реализации скачка тяжких преступлений уже созрели, но не хватает некоторого дополнительного спускового механизма. Таким механизмом может быть резкое изменение природных и социальных факторов, а также даже дополнительный прием алкоголя во время продолжительных праздников. В пользу важности роли спусковых механизмов свидетельствует определенная коррелируемость режима числа убийств с такими, несомненно эмоциогенными, а не рассудочными событиями, как случаи самоубийств и сопротивлений органам правопорядка.

Выявленные закономерности в режиме преступности и алгоритмы прогноза скачков числа тяжких преступлений могут быть использованы в дальнейшем как компоненты системы мониторинга уровня социальной напряженности в социуме.

Настоящая работа касалась вопросов «саморазвития» режима преступности, как одной из сторон поведения социума. Для создания расширенной системы мониторинга на основе выявляемых в рамках синергетики универсальных закономерностей поведения сложных нелинейных иерархических систем ее необходимо дополнить компонентами, учитывающими воздействие внешних социально-экономических, техногенных и природных факторов.

Литература

1. Николис Г., Пригожин И. Самоорганизация в неравновесных системах. – М.: Мир, 1979.
2. Хекен Г. Синергетика. Иерархия неустойчивостей в самоорганизующихся системах и устройствах. – М.: Мир, 1985.
3. Bak P., Tang C. Earthquake as a self-organized critical phenomenon// J. Geophys. Res. 1989. V.94. P.15635-15637.
4. Арнольд В.И. Теория катастроф. – М.: Наука, 1990.
5. Lorenz E.N. The essence of chaos. – London: U.C.L. Press Ltd., 1993.
6. Turcotte D.L. Chaos, fractals, nonlinear phenomena on Earth sciences// U.S. National Report to IUGG 1991-1994// Rev. of Geophys. supplement. AGU. 1995. P.341-343.
7. Keilis-Borok V.I. (eds.). Intermediate-term earthquake prediction: models, algorithms, worldwide tests// Phys. Earth Planet. Inter. 1990. V.61, N1-2 (Spec. Iss.).
8. Keilis-Borok V.I, Soloviev A. (eds.). *Nonlinear Dynamics of the Lithosphere and Earthquake Prediction*. Springer, Heidelberg, 1-36, 2003.
9. Lichtman A., Keilis-Borik V.I. Pattern recognition applied to presidential elections in the United States 1860-1980; Role of integral social, economic and political traits// Proc. Nat. Acad. Sci. USA. 1981 V.78. P.7230-7234.
10. Lichtman A., Keilis-Borik V.I. Aggregate-level analysis and prediction of mid-term senatorial election in the United States, 1974-1986// Proc. Nat. Acad. Sci. USA. 1989 V.86. P.10176-10180.
11. Keilis-Borok V., Stock J.H., Soloviev A., Mikhalev P. Pre-recession pattern of six economic indicators in the U.S.A.// Journal of Forecasting. 2000. V.19, P.65-80.
12. Гребенюк Е.А., Кузнецов И.В. Применение методов последовательного анализа для прогнозирования резких скачков случайных временных рядов// Автоматика и Телемеханика. 1997. Т.11. С.65-75.
13. Kuznetsov I, Grebenuk E, Muratov D. About forecasting of crisis events// Труды международной конференции «Математическое моделирование социальной и экономической динамики». Москва. 2004. С.174-177.
14. Александров С.И., Гамбурцев А.Г., Гамбурцева Н.Г., Иванов-Холодный Г.С., Корсак М.Н., Лагунов С.И., Новиков А.М., Новоселова О.А., Олейник О.В. Комплексный экологический мониторинг и атлас временных вариаций. Атлас временных вариаций природных, антропогенных и социальных процессов. Т.3, М., 2002, 160-169.
15. Киселев С.Л., Коимшиди Г.Ф., Гамбурцев А.Г. Динамика зарегистрированных преступлений в России за 1991-2001 гг. Атлас временных вариаций природных, антропогенных и социальных процессов. Т.3, М., 2002, 576-583.
16. Молчан Г.М. Оптимальные стратегии в прогнозе землетрясений. Современные методы интерпретации сейсмологических данных// Выч. сейсмология. №24. С.3-18. – М: Наука, 1991.
17. Родкин М.В. Модель развития синергетического эффекта при сильных катастрофах// Геоэкология, 2005, №1, 1-7.