

Варианты метода «Гусеница»-SSA для анализа многомерных временных рядов

Ссылка на полную версию работы находится на странице
<http://www.gistatgroup.com/gus/books.html>

4. Примеры обработки рядов методами SSA

В этом разделе мы (на примере двумерного ряда продаж вин в Австралии) кратко опишем технологию применения двумерного варианта MSSA. Методики применения одномерного SSA и комплексного CSSA аналогичны. В заключение на том же примере будут продемонстрированы различия между тремя модификациями SSA.

4.1. Методика применения MSSA

Рассмотрим два временных ряда продаж вин (FORT и DRY – крепленые и сухие вина) в Австралии, ежемесячно, с января 1980 года по ноябрь 1993 года ($N=167$), в тысячах литров [9]. Ряды изображены на рис.1 и рис.2. Видно, что ряд FORT имеет почти постоянный тренд, а ряд DRY – тренд, близкий к линейному и с небольшим углом наклона. Также видно, что в рядах присутствуют сезонные колебания, причем годовое поведение продаж вин имеет достаточно сложную форму, что соответствует разложению сезонных компонент рядов на (возможно, модулированные) гармоники с почти всеми частотами из набора ($1/12, 1/6=2/12, 1/4=3/12, 1/3=4/12, 1/2.4=5/12, 1/2=6/12$).

Посмотрим, к какому разложению приведет применение к этим рядам двумерного варианта метода MSSA. Для улучшения разделимости возьмем длину окна L кратной периоду сезонной компоненты, т. е. 12-ти. Кроме того, возьмем L примерно равным половине длины ряда: $L = 84$. Тогда K также равняется 84.

На рис. 3, 4 и 5 приведены логарифмы собственных чисел сингулярного разложения траекторной матрицы рядов для двумерного MSSA, а также соответствующие собственные и факторные вектора. Форма графика собственных чисел говорит о том, что, скорее всего, первые 12 собственных троек соответствуют сигналу, а остальные – шуму.

Глядя на собственные и факторные вектора (рис. 4 и 5), можно идентифицировать первую и шестую собственные тройки как соответствующие главным образом тренду. Напомним, что первая половина факторных векторов соответствует первому ряду, а вторая – соответственно, второму.

Видно, что тренд не до конца разделился с более высокочастотными составляющими ряда. Анализ показывает, что произошло небольшое перемешивание тренда с сезонными компонентами рядов.

Далее, предложение 2.2 говорит о том, что гармоническим составляющим ряда соответствуют собственные и факторные вектора, также являющиеся гармоническими (в случае приближенной разделимости – приближенно гармоническими) рядами. Анализ собственных и факторных векторов (см. рис. 4 и 5), показывает, что 2-я и 3-я собственные тройки соответствуют годовой периодичности (период 12), 4-я и 5-я – четырехмесячному периоду, 7-я и 8-я – полугодовой (период 6) периодичности, а частоты $1/2.4$ и $1/3$ смешались и содержатся в собственных тройках 9–12. Компонента периода 2 в данном разложении не выделилась.

Для нахождения пар собственных векторов, соответствующих одной гармонике, очень удобным оказывается попарное изображение собственных

векторов на двумерных скаттерграммах (рис.6). В силу ортогональности собственных векторов при таком изображении правильный (выпуклый или звездчатый) многоугольник, возможно, несколько искаженный, идентифицирует гармонику. Смещение гармоник может быть выявлено с помощью периодограммного анализа, а также на основе анализа собственных троек, имеющих близкие собственные числа.

На рис. 7 и 8 представлены первые шесть компонент разложения исходных рядов, полученные без группировки (т.е. операция диагонального осреднения применена непосредственно к каждой элементарной матрице). Рис. 7 соответствует разложению первого ряда (FORT), а рис. 8 – разложению второго ряда (DRY).

Такая «элементарная» группировка не приводит к разделимым, а, следовательно, интерпретируемым компонентам ряда. Для того чтобы согласованно разложить каждый из рядов на тренд, сезонную компоненту и шум, нужно сгруппировать собственные тройки согласно их порядковым номерам на 3 группы: $I_1=\{1,6\}$, $I_2=\{2-5,7-12\}$, I_3 – все остальные собственные тройки.

На рис. 9–11 приведено соответствующее разложение исходных рядов. Видно, что тренд (рис. 9) достаточно гладкий (т.е., смешение с сезонной компонентой незначительное; к тому же вклад шестой собственной тройки небольшой). С помощью рис. 10 можно сопоставить сезонное поведение продаж вин для рассматриваемых рядов.

Для более детального изучения сезонных компонент может оказаться полезным сопоставление их гармонических составляющих (рис. 12–14). На рис. 12 и 13 изображены годовые и полугодовые гармонические составляющие рядов. Их фазы смещены друг относительно друга примерно на четверть периода. На рис. 14 изображены суммы остальных выделенных гармоник. Видно, что они почти совпадают, как по фазе, так и по амплитуде. Чтобы проинтерпретировать такое соответствие, нужна дополнительная информация о специфике употребления и продаж вин в Австралии, а также о технологии сбора данных.

4.2. Сравнение с другими вариантами

Проведем сравнение результатов рассмотренного выше MSSA с CSSA и с применением одномерного SSA к каждому из рядов по отдельности. В таблице 1 приведены результаты идентификации первых четырнадцати собственных троек разложений траекторных матриц с $L = 84$. Числа в клетках таблицы означают периоды гармоник (возможно, модулированных), присутствующих в соответствующей собственной тройке. Жирным шрифтом выделены периоды тех гармоник, наличие которых существенно.

Таблица 1. Идентификация собственных троек для разложения SSA, MSSA и CSSA

	FORT, SSA	DRY, SSA	FORT+DRY, MSSA	FORT+DRY, CSSA
1	тренд	тренд	тренд+12	тренд
2	12	12	12	12
3	12	12	12	4
4	4	4	4	4
5	4	4	4	6
6	6	3+6+2.4	тренд+6+4+3+2.4	3+12+2.4
7	6	3+6	6	12+2.4+3
8	2.4	2.4+3+6	6	2.4+3+12
9	2.4	3+6	3+2.4	2.4
10	3	2.4+3+6	3+2.4	3
11	3	2.4	2.4+3	6
12	шум	тренд	2.4+3	тренд
13	шум	тренд	шум	шум
14	шум	2	шум	шум

Опишем результаты сравнения:

- в смысле слабой разделимости одномерный метод показал себя лучше: для ряда DRY выделился более подробный тренд и обнаружилась гармоника периода 2; однако еще раз обратим внимание, что при применении SSA к рядам по отдельности нужно проводить дополнительную работу по сопоставлению компонент рядов, имеющих одинаковый смысл.;
- в смысле сильной разделимости результат применения одномерного метода сильно зависит от ряда: для ряда FORT вообще не произошло перемешивания компонент, в то время как для ряда DRY сильно перемешались три гармонические компоненты;
- метод MSSA частично убирает перемешивание для ряда DRY, однако добавляет его в ряд FORT;
- в одномерных SSA разложениях, а также в MSSA, компоненты, соответствующие одной гармонической составляющей ряда, идут подряд, что облегчает их идентификацию;
- в CSSA такой порядок компонент нарушен, что согласуется с теорией: из-за сдвига по фазе годовых (и полугодовых) гармонических составляющих исходных рядов соответствующие собственные тройки расположены не подряд; это привело, например, к смешению 12-месячной гармоник с более высокочастотными (заметим, что для комплексного случая «потерянная» (неучтенная при восстановлении) компонента приводит к искажению фазы восстановленной компоненты ряда, в то время как для MSSA – к искажению амплитуды).

Таким образом, ни один из методов не приводит к полному разделению всех интерпретируемых компонент для обоих рядов. Если сравнивать MSSA с CSSA, то можно говорить о некотором преимуществе двумерного SSA над комплексным: единственными компонентами, которые существенно смешиваются MSSA, являются слабые гармоники с периодами 3 и 2.4, в то время как CSSA не позволяет «чисто» выделить основную 12-месячную гармонику.

В то же время, если бы амплитуды двух основных (12-месячной и 4-месячной) гармоник были бы почти одинаковы, то, возможно, результат сравнения MSSA с CSSA был бы противоположным.

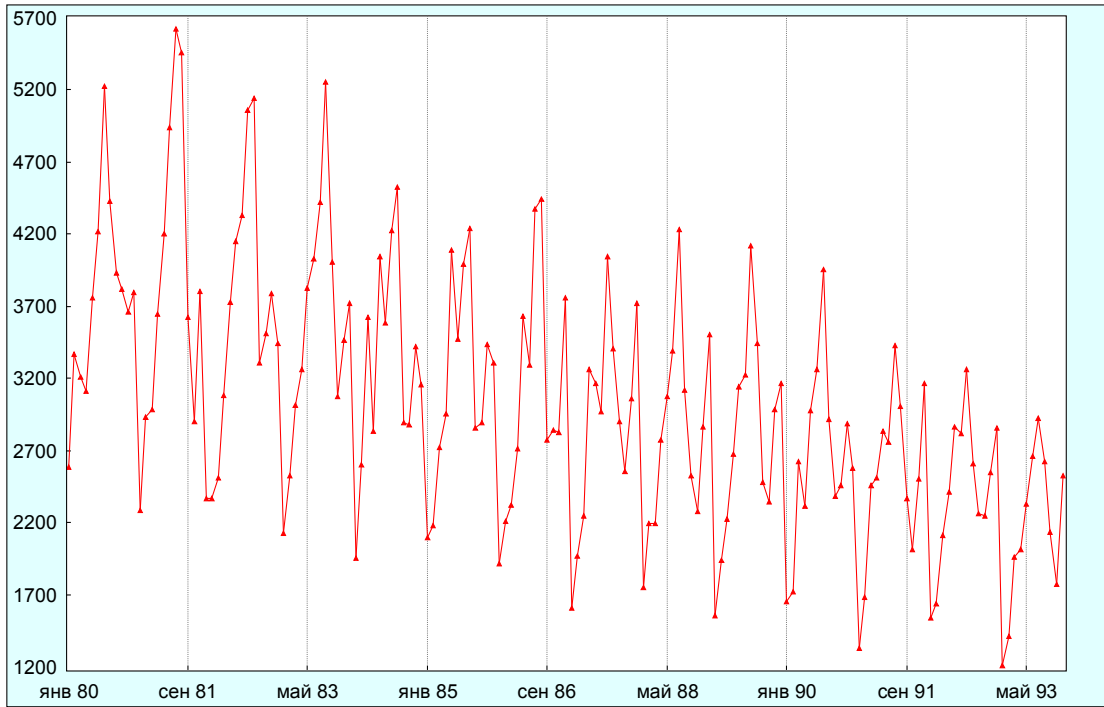


Рис. 1. Временной ряд FORT (объем продаж крепленых вин в Австралии)

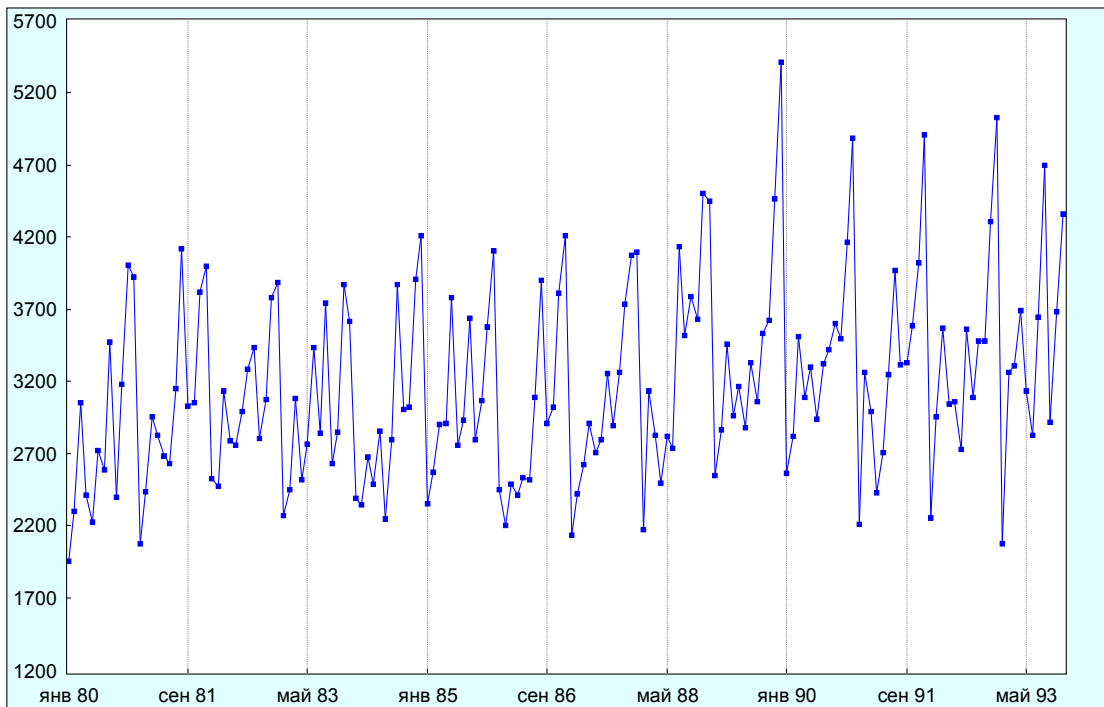


Рис. 2. Временной ряд DRY (объем продаж сухих вин в Австралии)

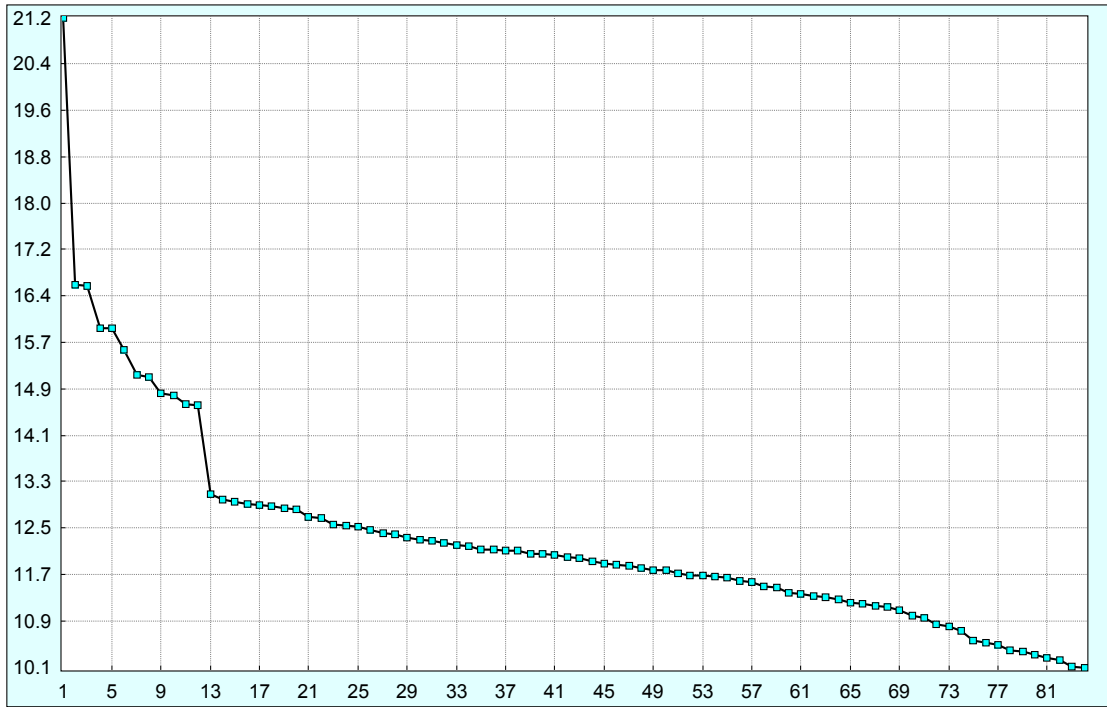


Рис. 3. MSA: логарифмы собственных чисел

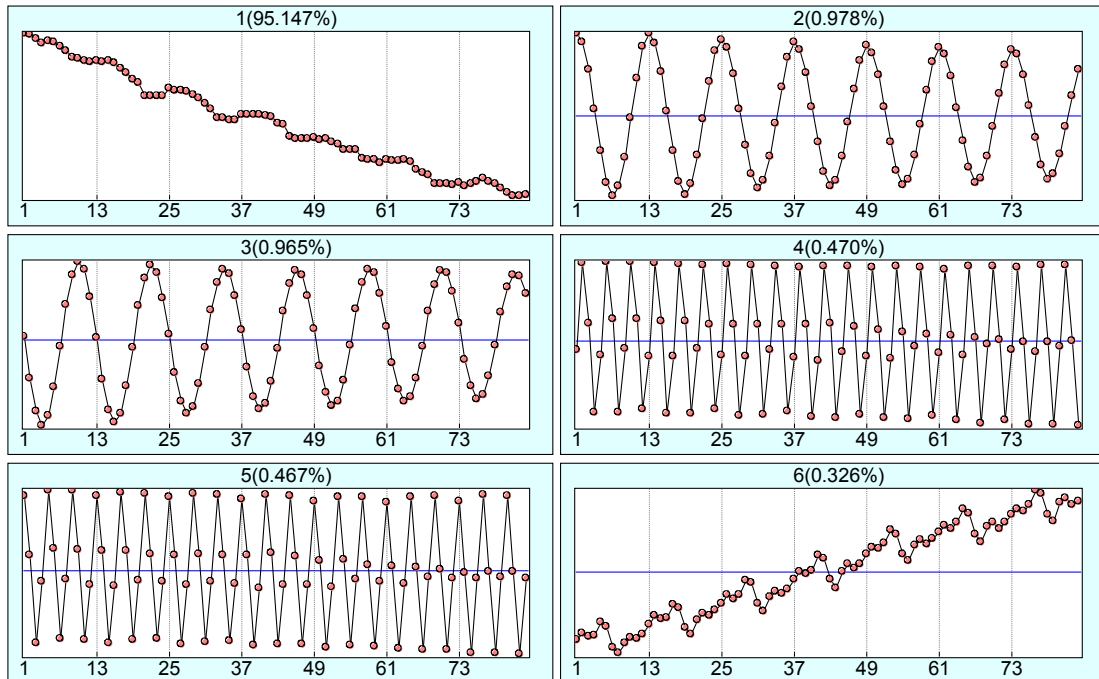


Рис. 4. MSA: собственные вектора

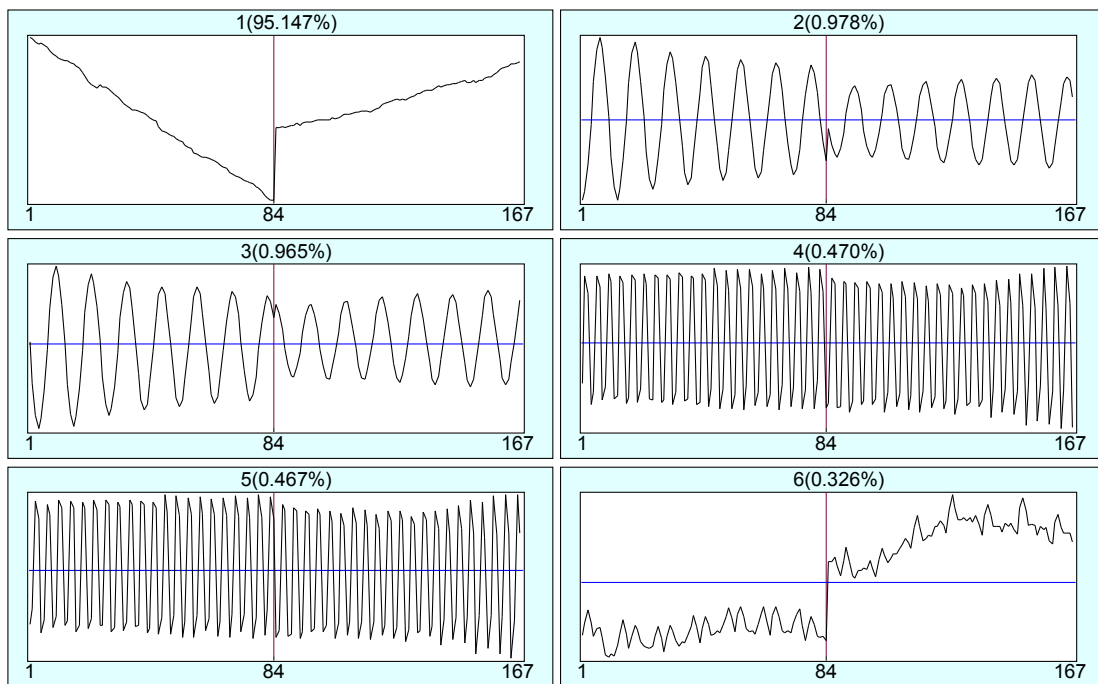


Рис. 5. MSA: факторные вектора

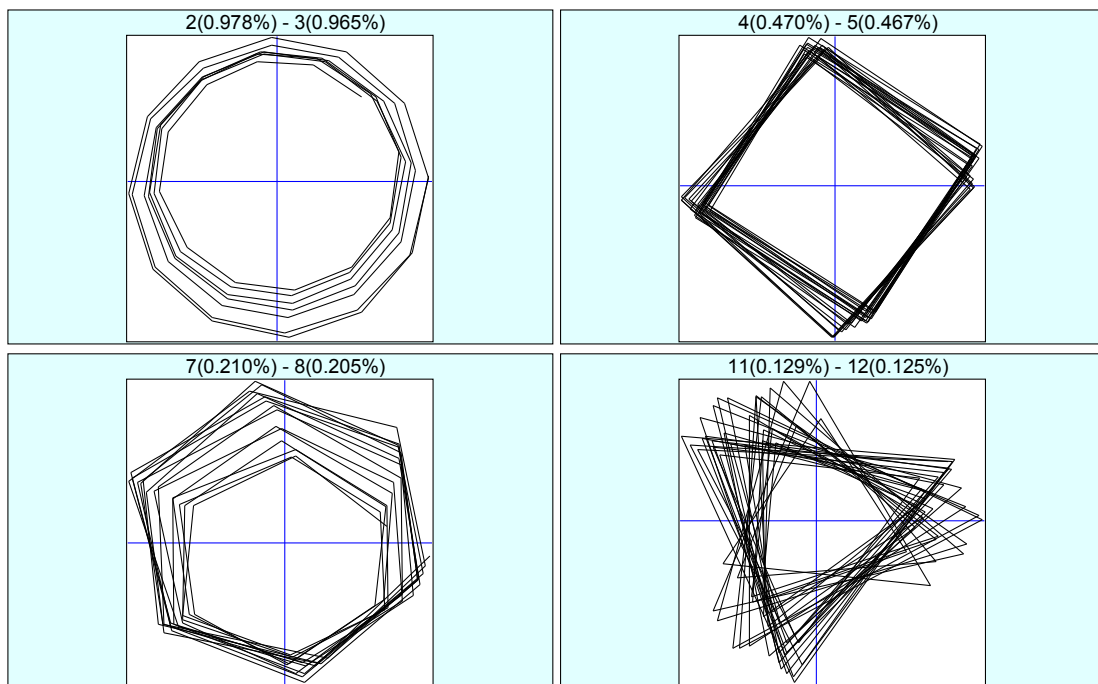


Рис. 6. MSA: двумерные скаттерграммы собственных векторов

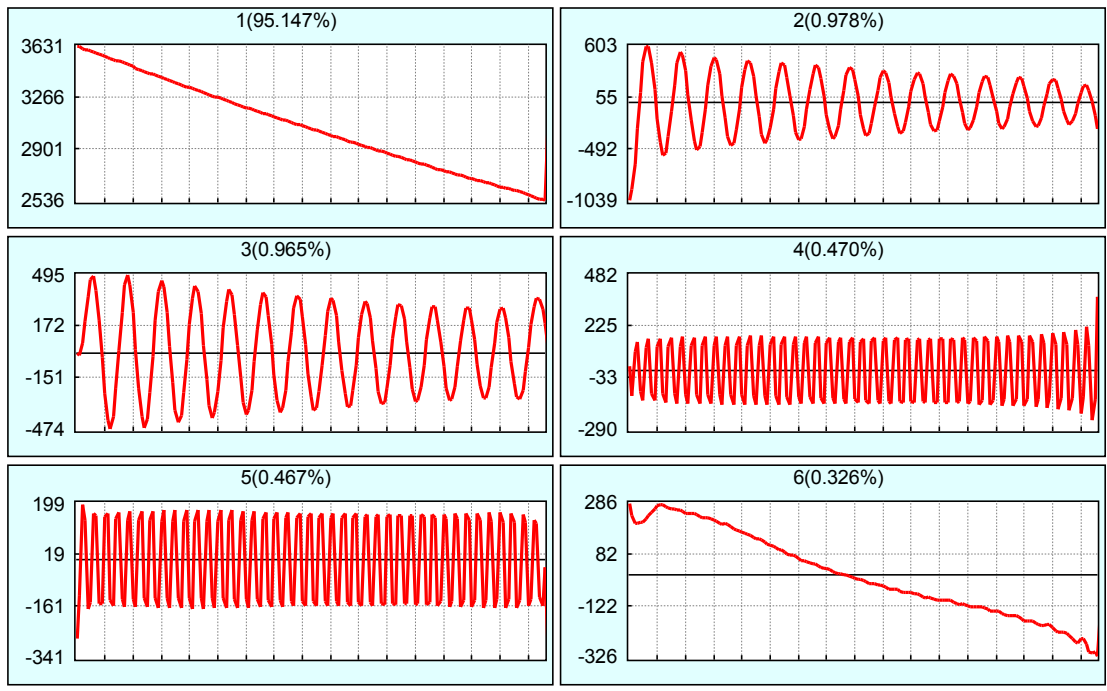


Рис. 7. MSSA-разложение ряда FORT

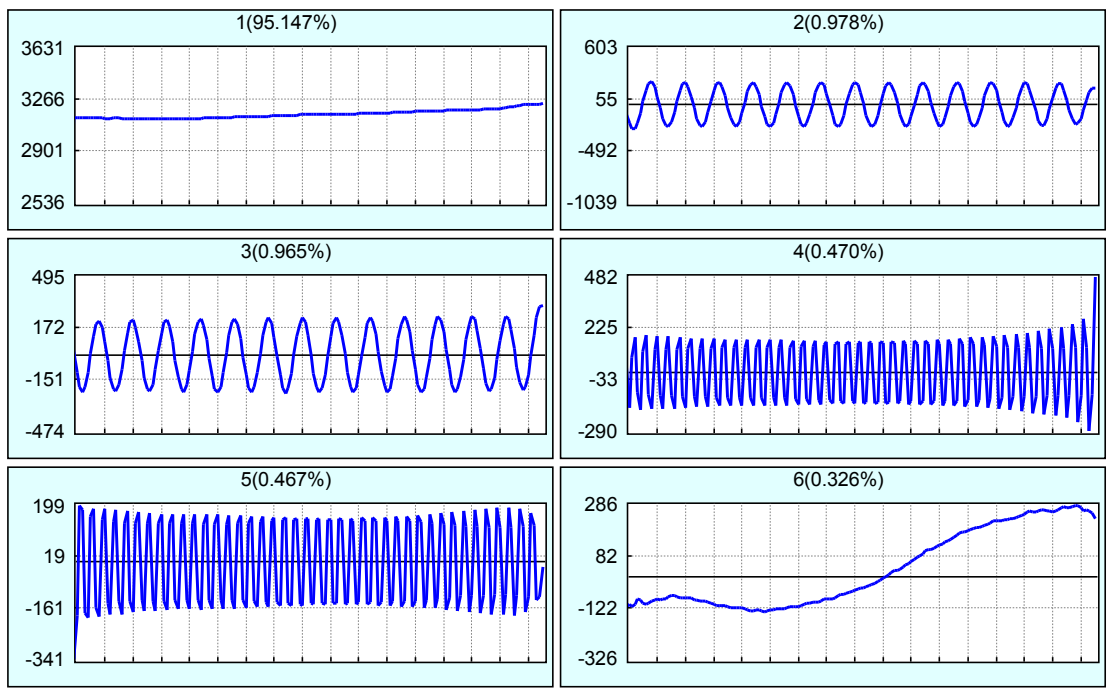


Рис. 8. MSSA-разложение ряда DRY

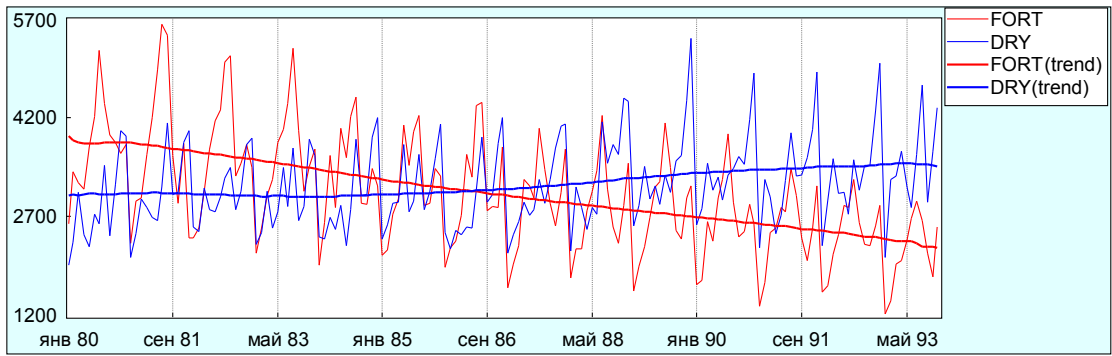


Рис. 9. MSA: тренды

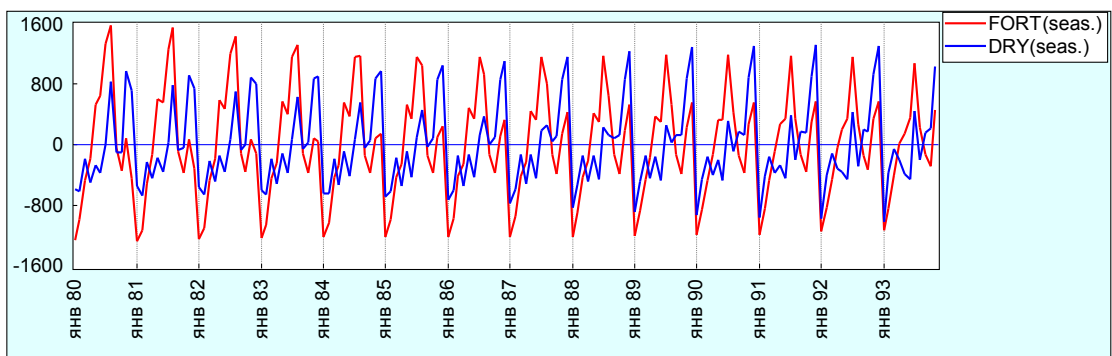


Рис. 10. MSA: сезонные компоненты

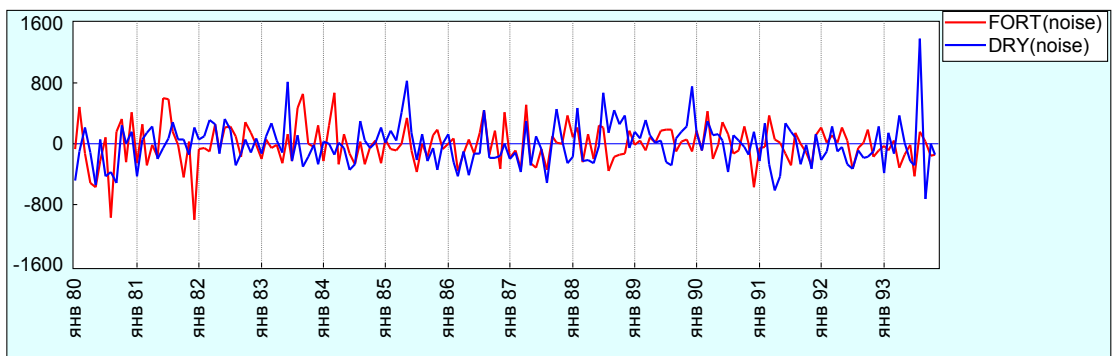


Рис. 11. MSA: шумовые компоненты

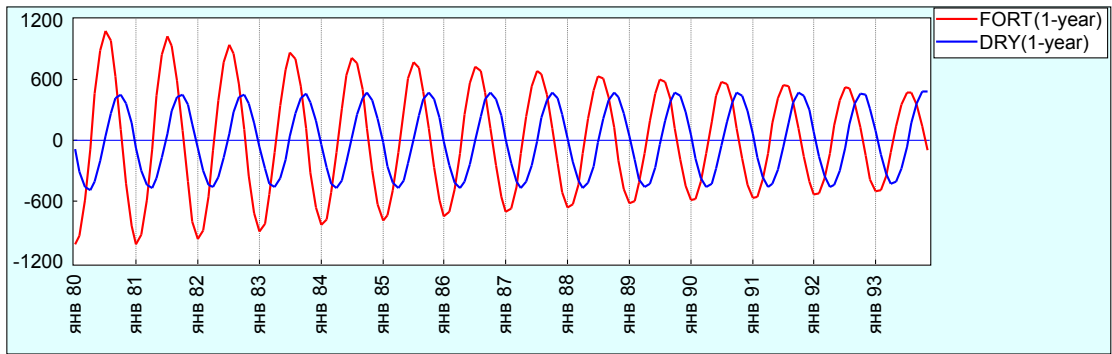


Рис. 12. MSSA: годовая периодичность, сдвиг на четверть периода

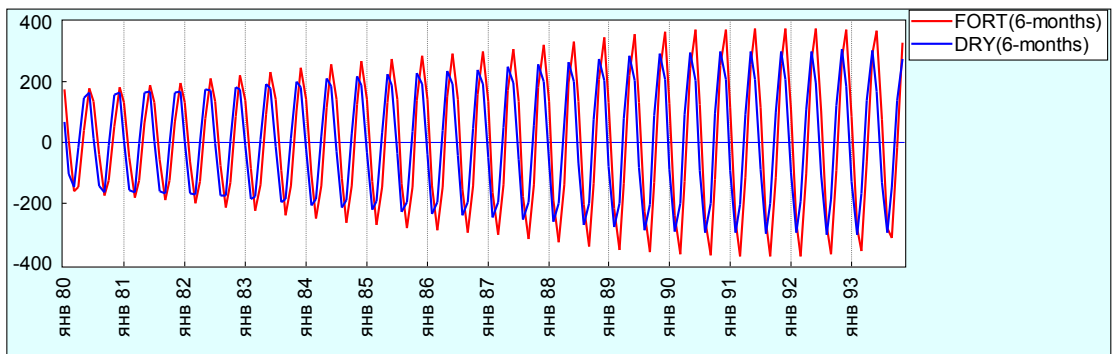


Рис. 13. MSSA: полугодовая периодичность, сдвиг на четверть периода

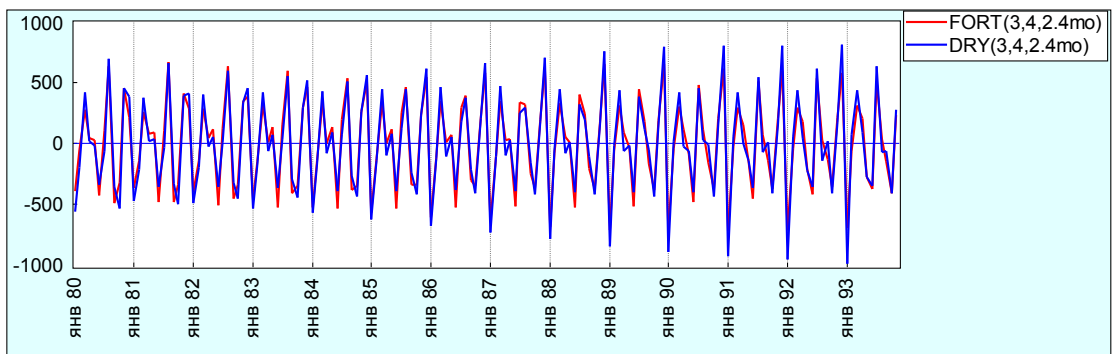


Рис. 14. MSSA: сумма 3, 4, 2.4 – месячных периодичностей, совпадение фаз