

ПРИМЕНЕНИЕ МЕТОДА САМООРГАНИЗУЮЩИХСЯ ОТБРАЖЕНИЙ ПРИЗНАКОВ К АНАЛИЗУ ЭКСТРЕМАЛЬНЫХ ОСАДКОВ

В.Л.Посошков, А.В.Прусов

Морской гидрофизический институт
НАН Украины
г. Севастополь, ул. Капитанская, 2
E-mail: oaoi@alpha.mhi.iuf.net

В работе проверяется один из возможных механизмов воздействия глобальных полей циркуляции и влаги на локальные осадки. Использовались многолетние наблюдения за суточными суммами зимних осадков в Севастополе. Применялась техника самоорганизующихся отображений признаков (СООП) и искусственная нейронная сеть (ИНС). Выделены образцы циркуляции, ответственные за события экстремальных осадков (СЭО).

Введение. Моделирование режима осадков при изучении изменчивости климата и для гидрологических приложений требует специального рассмотрения механизма осадкообразования в конкретном регионе на различных временных масштабах. Имеется большое число методологий, способных воспроизводить режим осадков. Некоторые исследования отмечают высокую эффективность моделей даунскейлинга, использующих нелинейные подходы, например, искусственные нейронные сети, по сравнению с линейными [1]. Другие не находят большого различия в этих методиках [2] или даже не видят никакого улучшения по сравнению с другими подходами [3]. В этой работе выбран метод [4] самоорганизующихся отображений признаков в сочетании с искусственной нейронной сетью прямого распространения [5].

С помощью СООП исследовалось влияние глобальной циркуляции и атмосферной влаги на события экстремальные зимних осадков в регионе Крыма. Кроме того, ИНС как диагностический инструмент, применялся для воспроизведения суточных сумм осадков в пункте исследования, в качестве которого был выбран Севастополь как ближайший к узловой сетке реанализа пункт Крыма, обеспеченный многолетними данными наблюдений.

Данные. В качестве предикторных переменных в работе использовались двух-

срочные данные из глобального проекта реанализа NCEP/NCAR (National Centers for Environmental Prediction/National Center for Atmospheric Research), за 14 зим (декабрь-февраль) 1982-96 г.г... Целевой переменной (предиктантом) в анализе выбрана суточная сумма осадков из многолетнего массива наблюдений в Севастополе. Все двухсроковые переменные предварительно переводились в среднесуточные.

Были выбраны следующие параметры атмосферы: полусуточные значения удельной влажности на высоте 2м (q_2), полусуточные значения высоты геопотенциала 500 гПа (z_5), и толщины слоя 500-1000 гПа (Th_{5_1}), полусуточные значения относительной влажности на уровне 500 гПа (r_5) на глобальной сетке $2.5^\circ \times 2.5^\circ$.

Для переменных q_2 и r_5 информация извлекалась только в целевой координатной точке, а для переменных z_5 и Th_{5_1} - еще и в восьми ближайших к ней узлах. Все переменные сдвигались по времени на 24 и 48 часов. В итоге обрабатывалось $m=1260$ (14×90) векторов размерности 60 ($1 \times 3 + 1 \times 3 + 1 \times 3 \times 9 + 1 \times 3 \times 9$).

Методология нейронной сети. Применяемая методология является диагностическим инструментом для выявления аномальных атмосферных мод, ответственных за события экстремальных осадков. Она состоит из трех основных шагов.

- Разбиение атмосферного воздействия на классы, или климатические моды (типы погоды).

- Выявление крупномасштабных климатических аномалий, ассоциированных с СЭО.

- Получение эмпирических передаточных функций между атмосферным воздействием и локальными осадками.

Для обеспечения первого пункта использовалась техника СООП Кохонена [4]. Как классификатор признаков, она аналогична кластерному анализу, но имеет то преимущество, что генерирует нелинейную классификацию. Процесс, с помощью которого образуется СООП, является алгоритмом обучения без учителя. Его можно использовать для нахождения кластеров подобных событий во входном потоке данных и идентифицировать некоторые лежащие в

его основе динамические структуры. Техника СООП относительно нова в исследованиях климата. Выбор числа классов является произвольным. Когда используется небольшое число СООП-образов, классификация событий обладает чрезмерной общностью. Это означает, что только первые моды атмосферы будут идентифицированы, а непрерывность атмосферных событий будет утрачена. Вместе с тем важно трактовать состояние атмосферы как непрерывно меняющуюся функцию, а не как ограниченный ряд дискретных состояний [6]. В этом исследовании было выбрано 6 классов, что означает карту изображений размерностью 3×2 ; при этом в среднем приходится 210 суток на класс.

Центральным результатом самоорганизации является *нелинейная проекция функции плотности вероятности многомерного входного вектора на двумерную карту*. Ожидается, что по окончании тренировочного процесса лучшая карта (лучшее отображение) дает наименьшую среднеквадратическую ошибку.

По окончании обучения алгоритма СООП временные ряды компонент атмосферного контроля конструируются специальным образом под СООП-классы. Затем находятся средние суточные суммы осадков (из фактического ряда наблюдений в Севастополе), ассоциированных с каждым классом табл. 1.

Для всего нормированного ряда осадков средняя величина=0.60, а ср.кв.откл.=1.0. Если упорядочить номера классов по уров-

ню среднего, то номера классов в упорядоченной группе должны идти в последовательности: 1, 3, 4, 2, 5, 6.

Таблица 1 – Статистика наблюдений за осадками в пределах каждого класса

Номер класса	Среднее	Ср. кв. отклонение	Число дней
1	0.20	0.47	319
2	0.68	1.30	170
3	0.49	0.80	209
4	0.67	1.10	172
5	0.74	1.08	134
6	1.01	1.10	254

Это позволяет ввести упорядоченную категориальную переменную, скажем S_{CAT} , отвечающему определенному классу. При этом проводилось не линейное упорядочивание, как в работе [5], а каждому классу присваивалась величина, пропорциональная среднесуточной норме осадков, соответствующей данному классу. Категориальная переменная S_{CAT} добавлялась к временным рядам атмосферного воздействия, чтобы выделить каждый день в соответствии с каждым из шести классов. Это образует в сумме 61 предикторную переменную - они используются затем в ИНС прямого распространения для генерирования суточных сумм осадков P в узловой точке:

$$P = f(_m \bar{X}_{61}) = f(q2_3, r5_3, Th5_1_{27}, z5_{27}, S_{CAT}). \quad (1)$$

В последние годы применение ИНС быстро распространилось среди приложений климата. ИНС дает альтернативу статистическим моделям, которые часто являются ограниченными из-за предположений о линейности, нормальности и т.д. Использовалась версия ИНС, как в работе [7]. В результате получались передаточные функции (диагностические нелинейные связи), которые могут быть использованы в дальнейшем для моделирования атмосферных осадков.

Кроме того, ИНС применялась для воспроизведения суточных норм осадков от-

дельно для каждой климатической моды. Композитные временные ряды атмосферных предикторов каждого образца СООП подавались на вход в нейронную сеть через входной слой (N_i), который связан со скрытым слоем системой взвешенных связей (N_h). Скрытый слой связан с единственным выходом ($N_o=1$), этот выходной сигнал сравнивался с фактическими осадками в течение тренировочного процесса. Число свободных параметров равно

$$W = (N_i + 1)N_h + (N_h + 1)N_o \quad (2)$$

Выбор контролирующих переменных. Средние поля крупномасштабной циркуляции отвечают за динамику циркуляции, но они не в состоянии уловить среди других локальных воздействий процессы, связанные с термодинамикой и водяным паром. В данном анализе выбор (селекция) предикторов для осадкообразования на суточном масштабе основан на простом рациональном принципе: динамическая, термическая и влажная компоненты. Крупномасштабная циркуляция частично улавливается полями z5 и Th5_1, а также индексом СООП. Бароклиновые системы, такие как фронтальные и циклонические возмущения ассоциируются с температурными градиентами атмосферного слоя; таким образом, использована толщина слоя 500-1000 гПа. Наконец, удельная влажность на уровне 2м и относительная влажность на уровне геопотенциальной поверхности 500 гПа применялись, чтобы принять в расчет влагу и конвекцию. Временные сдвиги в переменных атмосферного воздействия использовались с тем, чтобы переносить информацию от крупномасштабного состояния к локальному.

Результаты СООП классификации. Цель этого раздела – дополнить наше понимание физики атмосферных процессов, связанных с СЭО в Крыму (на основе массива данных по Севастополю). Это делалось путем изучения характеристик событий экстремальных аномалий климата с помощью СООП-классификации, которая применялась для конструирования рядов осадков, соответствующих каждой климатической моде. Классы, относящиеся к экстремально обильным и экстремально малым осадкам для ближайшей к Севастополю точки координатной сетки реанализа (45° с.ш., 35° в.д.) использовались для составления образцов (композитов) крупномасштабных полей Z5 и q5. Эти средние поля и их аномалии были отобраны для исследования физических характеристик атмосферы во время СЭО в пункте исследования.

a) Аномалии циркуляции. На рис.1а,б показана высота геопотенциала 500 гПа для экстремально дождливых и экстремально засушливых условий, соответственно. Рис.1а показывает, что аномально влажные события в Крыму связаны с сильной среднетропосферной циркуляцией над центральной и восточной Европой. Расщепление струи над субтропической Атлантикой и ложбина над Средиземным морем предполагают область сильного входа субтропической струи над северной Африкой и восточным Средиземноморьем в течение синоптических условий, приводящих к влажным событиям. В противоположность этому, аномально сухие события сопровождаются усиленной зональной циркуляцией над восточной северной Атлантикой и Европой, а также гребнем, смещенным к восточной Европе (рис. 1 б). Аномалии циркуляции, относящиеся к дождливым и засушливым событиям, в центральной Европе почти противоположны друг другу (рис.1 в,г). Во время дождей возникает система низкого давления над восточной Европой, ограниченной Азорским антициклоном и холодным Сибирским антициклоном, и среднетропосферный меридиональный поток (рис.1а) способствует адвекции холодного воздуха в направлении центрального и восточного Средиземноморья, где взаимодействие с теплыми морями генерирует сильную циклоническую активность. Относительно низкое давление над Крымом зимой связано с осью максимума частоты холодных фронтов в центральном бассейне Средиземного моря, которая формируется южнее Италии, откуда депрессии направляются в сторону Черного моря и юга Украины. И наоборот, засуха характеризуется расширением Азорского максимума до Южной Европы и Средиземного моря. Относительно сильный среднетропосферный гребень (рис.1б) и антициклон во время засухи, указывают на смещение Северо-Атлантической струи на север. Таким образом, похоже, что СЭО регулируются изменениями местоположения струйного течения и сопутствующими шторм-треками.

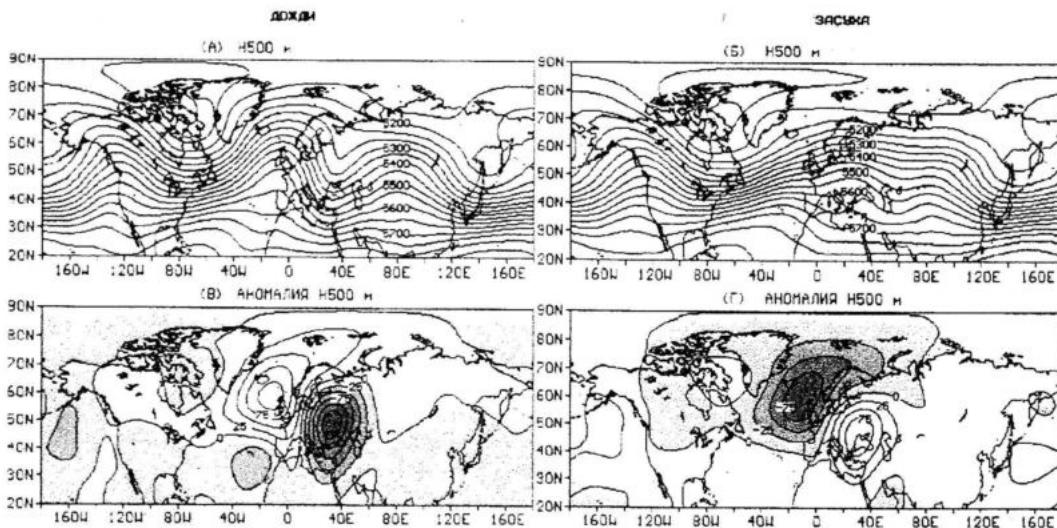


Рис. 1 – Композитные карты средней высоты (а, б) и аномалий (в, г) геопотенциальной поверхности 500 гПа (z_5), соответствующие двум экстремальным случаям влажного (а, в) и сухого (б, г) СООП классов. Изолинии проведены через 25 м

б) Роль среднетропосферной влаги во время экстремальных событий. Как указывалось выше, среднетропосферная ложбина в центральной и восточной Европе благоприятствует адвекции холодного воздуха к восточному Средиземноморью, где неустойчивость пограничного слоя способствует циклонической активности и переносу влаги к Восточной Европе. Изменения в среднем и переменном потоках воздейстуют на перенос и конвергенцию влаги и, следовательно, могут быть связаны с изменчивостью региональных осадков.

Анализ нескольких синоптических случаев, принадлежащих влажным и сухим СООП-композитам, показал подобие полей q_2 влажности на поверхности в бассейне восточного Средиземноморья и над областью исследования как во время влажных, так и сухих событий. Однако, наблюдаются существенные различия в образцах относительной влажности на уровне геопотенциальной поверхности 500 гПа (г5). Хотя влажность у поверхности земли может играть важную роль в дестабилизации пограничного слоя, на более высоких уровнях она может обладать более определяющим фактором во время крупных событий осадков. Похоже, для того, чтобы нарушить нисходящую циркуляцию, типичную для сухих событий, необходимо не только термическое воздействие на нижних уровнях (т.е. пар), но также и на верхних уровнях вкупе с динамическими управляющими

процессами, чтобы поднять влагу. Меридиональное течение и конвергенция от поверхности к средней тропосфере, а также язык влаги на уровне z_5 , который advectируется с восточного Средиземноморья и Черного моря в изучаемую область, накладывают различие между влажными и сухими синоптическими событиями осадков.

Генерирование суточных сумм осадков с помощью ИНС. В последних двух частях мы демонстрировали полезность СООП классификации для понимания изменчивости СЭО на синоптическом масштабе. Затем при помощи ИНС проводилось моделирование суточных норм осадков. При этом использовались те же предикторные переменные, что и при СООП классификации. Генерирование осадков проводилось в двух вариантах.

В первом случае ИНС настраивалась в каждой из шести композитных выборок данных. Результаты содержатся в табл.2. В последней колонке таблицы приводится число Хансена-Купера [8] НК, которое мы использовали как числовую меру качества моделирования последовательности событий, относящихся к двум взаимоисключающим категориям. Идеальное моделирование соответствует случаю НК=1 (Более подробно об этом числе в работе [7]). В нашем случае за событие “дождь” принималась ситуация, когда суточная сумма осад-

ков превышала среднюю норму осадков в соответствующей выборке.

В другом случае использовалась категориальная переменная S_{CAT} , которая добавлялась к предикторным переменным. Заметим, что использование этой переменной не привело к улучшению результатов по сравнению с содержащимися в табл.2. Результаты ИНС показали, что из атмосферных параметров, которые были использованы, наиболее важными предикторами осадков на суточном масштабе оказались $Th5_1$, $z5$, $r5$.

Таблица 2 - Числовые характеристики качества моделирования при помощи ИНС в пределах каждого СООП класса

Номер класса	Среднее набл.	Среднее ИНС	НК
1	0.20	0.26	0.40
2	0.66	0.85	0.66
3	0.50	0.42	0.47
4	0.63	0.59	0.59
5	0.88	0.80	0.48
6	1.02	0.70	0.34

Из табл.2 видно, что ИНС завышает среднюю норму осадков в сухом классе и наоборот, занижает во влажном. Этот эффект представляет собой типичную проблему даунскейлинга осадков. События экстремально больших осадков трудно моделировать не только ввиду сложности и изменчивости управляющих ими физических процессов, но также из-за редкости этих событий на суточном временном масштабе. Невысокие значения числа НК в табл.2 свидетельствуют о трудности моделирования осадков на суточном масштабе с помощью ИНС. Использование слишком большого числа предикторов при сравнительно небольшой выборке данных приводит к тому, что число искомых весовых параметров превышает число наблюдений в выборке данных. Так, в нашем случае согласно формуле (2) при использовании 5-ти нейронов в скрытом слое число параметров разрастается до величины $(61+1)\times 5 + (5+1)\times 1 = 316$, что существенно превышает среднюю длину выборки, равную 210. Это приводит к перенасыщению сети параметрами на стадии обучения и к утрате сетью

способности к обобщению на независимую выборку данных.

Таким образом, если в модели присутствует тщательно подобранный ряд предикторов, проблема адекватного воспроизведения событий осадков состоит не только в недостатке в модели других локальных функций влияния (процессы на поверхности суши, конвекция и т.п.), но и в недостаточной репрезентативности данных и большой изменчивости этих событий, которая во многих случаях обусловлена случайными процессами.

Заключение и выводы. В статье проверяются локальные и удаленные источники изменчивости, связанные с зимними осадками в регионе Крыма. При этом акцент делается на экстремальных событиях. За счет применения техники нелинейной СООП-классификации и ИНС прямого распространения продемонстрирована полезность СООП-ИНС подхода для идентификации пространственной и временной атмосферных структур, дающих существенный вклад в изменчивость осадков. Некоторые из характеристик климатических аномалий, полученные с помощью СООП, согласуются с прежними работами, основанными на технике эмпирических ортогональных функций и классического канонического анализа. В этой статье сделана попытка определить степень воздействия циркуляции, термической и влажной компонент на зимние осадки. Климатические аномалии, полученные за счет СООП классификации, обнаружили, что события экстремальных зимних осадков в Крыму на синоптическом масштабе в основном регулируются двумя факторами: изменениями местоположения струйного течения и сопутствующих шторм-треков, а также вариациями среднетропосферной влаги над восточным Средиземноморьем. Синоптические возмущения на нижних уровнях возникают, когда происходит незначительное ослабление крупномасштабного среднетропосферного зонального потока.

Несмотря на субъективизм в выборе числа мод СООП, это исследование демонстрирует, что климатические СООП аномалии, типичные для экстремальных условий, физически согласуются с локальным и удаленным воздействиями. В перспективе техника СООП может дать преимущество в

выявлении множественных источников изменчивости на различных временных масштабах. Даже 5-дневные и 10-дневные средние приводят к смешению экстремальных с неэкстремальными условиями и, таким образом, могут сгладить атмосферный отклик. Вместо этого СООП классифицирует дни в нелинейной форме, максимизируя различные атмосферные условия в каждой mode. Затем возможно "прокрутить" последовательность синоптических картинок с временными шаблонами различной длины в соответствии с состоянием атмосферы в конкретный сезон, что может помочь обнаружить развитие и возможную причину экстремальных событий.

Результаты ИНС показали, что среди использованных предикторов при воспроизведении суточных норм осадков наиболее значимыми оказались Th5_1, z5 и r5. Интересно, что переменная S_{CAT} не играет существенной роли в объяснении дисперсии суточных сумм осадков. Похоже, что толщина тропосфера, а также среднетропосферная циркуляция и влага несут много больше информации, чем простая категориальная переменная, связанная с климатической моделью. Модель ИНС в состоянии уловить синоптические события и хорошо отделяет сухой от влажного периода, но имеет тенденцию завышать (занижать) малые (большие) события. Данный анализ представляется полезным для диагностического исследования климата, в том числе даунскейлинга, а также для региональных и водоресурсных приложений.

ЛИТЕРАТУРА

1. Weichert, A., and G. Burger. Linear versus non-linear techniques in downscaling. Climate Res., vol 10, 1998, 83-93.

2. Wincler, J.A., J.P. Palutikof, J.A.Andresen, and C.M.Goodess. The simulation of daily temperature time series from GSM output. Part II. Sensitivity analysis of an empirical transfer function methodology. J. Climate, vol.10, 1997, 2514-2535.
3. Wilby, R.L., T.M.L. Wigley, D.Conway, P.D. Jones, B.C.Hewitson, J.Main, and D.S.Wilks. Statistical downscaling of general circulation model output.: A comparison of methods. Water Resour. Res. Vol. 34, 1998, 2995-3008.
4. Kohonen, T. Self-organized formation of topologically correct features maps. Bio. Cyber., vol 43, 1982, 59-69.
7. Cavazos, T. Large-scale circulation anomalies conducive to extreme events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. J. Climate, vol 12, 1999, 1506-1523.
6. Hewitson, B.C. Deriving regional climate change scenarios from GSMS. Water Research Comission Report K5/751, Pretoria, South Africa, 1998, 19 pp. {Available online at <http://www.wrc.org.za/>}
7. Посошков В.Л. Использование искусственных нейронных сетей для статистического оценивания ряда среднесуточных осадков. – Системы контроля окружающей среды, Севастополь, МГИ, 2002г., с146-152.
8. Hansen,A.W., and W.J.A. Kuipers. On the relationship between the frequency of rain and various meteorological parameters. //Koninklijk Nederlands Meteorologisch Institut, Meded.Verhand.-1965.- 81.- pp.2-15.
9. Посошков В.Л. Некоторые аспекты оптимального оценивания качества прогноза дискретных процессов. – Системы контроля окружающей среды, Севастополь, МГИ, 2002г., с153-158.