

ПЕРВЫЕ РЕЗУЛЬТАТЫ ОЦЕНКИ УСПЕШНОСТИ МЕЗОМАСШТАБНЫХ ЧИСЛЕННЫХ ПРОГНОЗОВ COSMO-RU, ВЫПУСКАЕМЫХ В РАМКАХ МЕТЕООБЕСПЕЧЕНИЯ ОЛИМПИАДЫ СОЧИ-2014

*А.Ю. Бундель, А.А. Кирсанов, А.В. Муравьев, Г.С. Ривин,
И.А. Розинкина, Д.В. Блинов*

*Гидрометеорологический научно-исследовательский центр
Российской Федерации
a.bundel@gmail.com, muravev@mecom.ru, Gdaly.Rivin@mecom.ru,
inna@mecom.ru, denisblinov@yandex.ru*

Введение

Оценка качества (верификация) прогнозов погоды является важнейшей составляющей метеорологического обеспечения зимних Олимпийских игр Сочи-2014. Сочинский регион имеет сложный рельеф с большими перепадами высот (отроги Большого Кавказского хребта), рядом находится незамерзающее Черное море. Для проведения соревнований требуется особая метеорологическая информация с большим временным и пространственным разрешением. Эти факторы делают особенно трудным как само прогнозирование, так и верификацию численных прогнозов, определяют необходимость разработки новых подходов к верификации. Для надежной верификации необходима густая сеть данных наблюдений. В районе проведения Игр были введены в действие дополнительные автоматические наблюдательные станции, однако их все равно оказалось недостаточно для применения новых, так называемых пространственных методов мезомасштабной верификации, как, например, верификация в окрестностях (neighborhood verification).

Для статистически значимых выводов необходимы длинные ряды данных прогнозов и наблюдений, а некоторые модели стали доступны лишь недавно. Так, например, непосредственно перед началом Игр для Сочинского региона была внедрена версия мезомасштабной модели COSMO-Ru1 с шагом сетки 1,1 км помимо уже имеющихся версий с шагами сетки 7 и 2,2 км.

В настоящей статье приведены лишь первые оценки качества прогнозов на период Игр. Кроме того, даны некоторые результаты по использованию новых характеристик качества прогнозов за зиму 2012/2013 гг.

Основное внимание посвящено модели COSMO-Ru [3], однако для сравнения будут привлекаться результаты модели NMMB. Большая часть оценок была построена в VERSUS (VERification System Unified Survey) – официальной программе верификации стран-членов консорциума COSMO.

Опыт Олимпиады: что оказалось полезнее всего, чего не хватало

Период Игр был аномально теплым, и все модели прогнозировали более низкие значения приземной температуры воздуха. Например, за первую неделю Олимпиады все модели прогнозировали температуру порядка +5 °С в Красной Поляне, в то время как в реальности наблюдалась температура около +15 °С. На этом фоне модель HARMONIE в этот период давала несколько лучшие прогнозы приземной температуры вплоть до высот ~1200–1400 м. Однако выше 1400 м прогнозы HARMONIE оказались менее точными.

Следует отметить, что в рамках проекта FROST-2014 (Forecast and Research in the Olympic Sochi Testbed) для синоптиков стал доступен широкий спектр прогностической продукции разных моделей. Очень полезными, по отзывам синоптиков, оказались прогнозы по модели COSMO-Ru1 [4], использующей рельеф ASTER с разрешением 30 м, в частности, комбинированные карты направления ветра и относительной влажности, по которым можно предсказывать движение линз влажного или сухого воздуха. В то же время традиционные точечные оценки таких прогностических

продуктов моделей высокого разрешения не являются показательными. Для всесторонней оценки таких полей требуются пространственные методы [7].

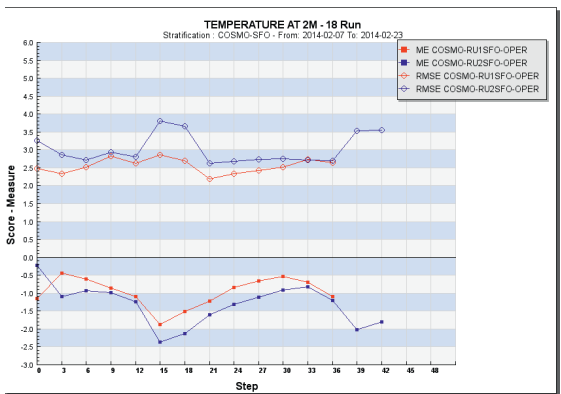
В настоящее время в Гидрометцентре России ведутся работы по созданию совместного анализа радарных и наземных наблюдательных станций, который сделает возможным применение данных методов. В следующем разделе приводятся некоторые традиционные оценки качества основных метеоэлементов.

Оценки качества основных метеоэлементов за период Олимпиады Сочи-2014 (февраль 2014 года)

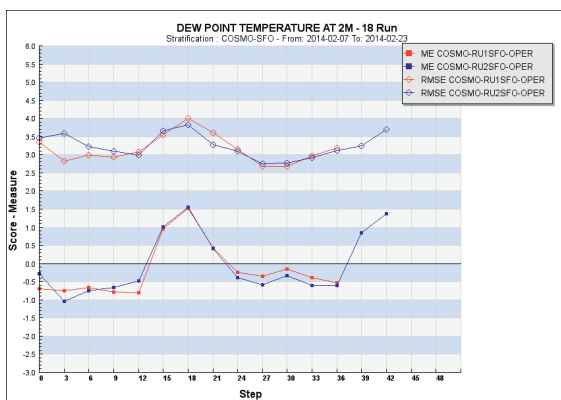
На рис. 1 приведены графики средней ошибки (ME) и среднеквадратической ошибки (RMSE) для моделей COSMO-Ru1 и COSMO-Ru2 за период Олимпиады (02–23 февраля 2014 года) для области интегрирования COSMO-Ru1, целиком находящейся в области COSMO-Ru2. Для всех элементов, кроме общей облачности, для сравнения бралась ближайшая модельная точка для каждой станции. Для общей облачности модельные точки осреднялись в радиусе 30 км вокруг станции. Для температуры воздуха на уровне 2 м (рис. 1а) отмечается некоторое преимущество прогнозов по модели более высокого разрешения COSMO-Ru1. Для скорости ветра оценка RMSE одинакова для обеих моделей, за исключением скачка в начальный момент интегрирования, который отсутствует в COSMO-Ru1. В COSMO-Ru1 практически отсутствует завышение скорости ветра, имеющееся в COSMO-Ru2 (рис. 1в). Для температуры точки росы (т. е. влажности воздуха) значимые различия оценок отсутствуют. Оценки общей облачности показывают попеременный успех то одной, то другой модели на протяжении периода прогноза. Отметим, что для общей облачности также не вполне показательны традиционные оценки. Кроме того, облачность наблюдалась на меньшем количестве станций, таким образом, статистическая значимость оценок невелика.

Особый интерес представляет верификация осадков. Данные модели и наблюдений такого высокого пространственно-временного разрешения получены в России впервые. На рис. 2 показан

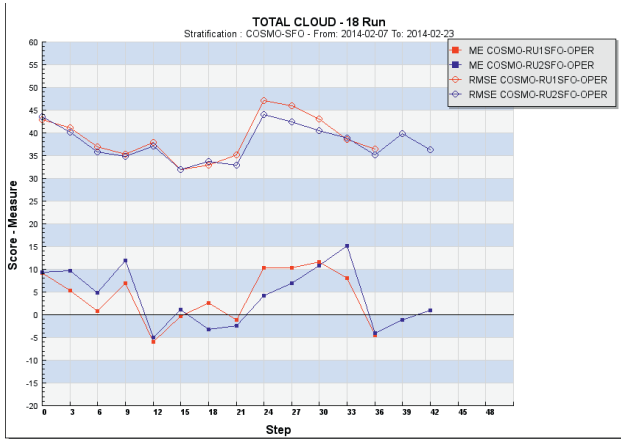
ход критерия Пирси-Обухова для 3-часовых осадков, превышающих порог 0,01 мм/3ч, т. е. наличия осадков (рис. 2а), и 1 мм/3ч (рис. 2б), в зависимости от заблаговременности для февраля 2014 г. для всего Сочинского региона с 3-часовым интервалом для моделей COSMO-Ru1 и COSMO-Ru2 для двух методов: ближайшая к станции модельная точка и осреднение модельных данных в радиусе 15 км вокруг станции.



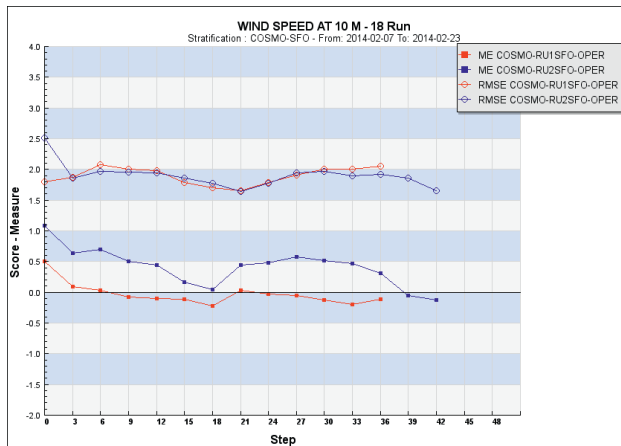
а)



б)

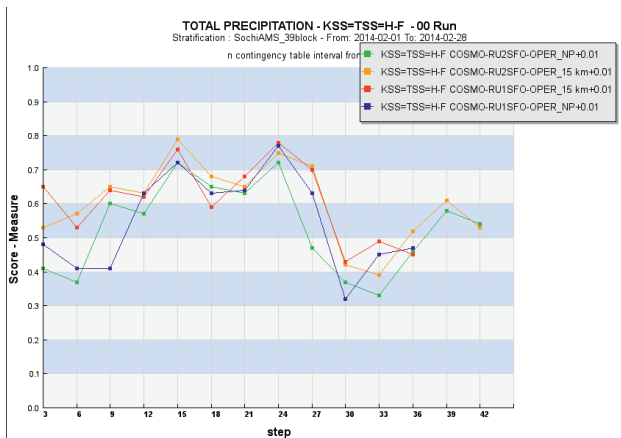


б)

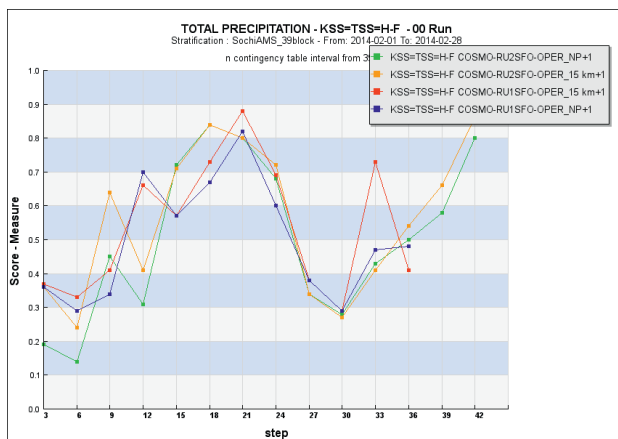


г)

Рис. 1. Графики средней ошибки (МЕ) и среднеквадратической ошибки (RMSE) для моделей COSMO-Ru1 (красные линии) и COSMO-Ru2 (синие линии) за период Олимпиады (02–23 февраля 2014 г.) для области интегрирования COSMO-Ru1: температура поверхности на уровне 2 м (а), точка росы (б), скорость ветра (в), общая облачность (г). Начальный срок 06 ч ВСВ.



a)



b)

Рис. 2. Критерий Пирси-Обухова для 3-часовых осадков, превышающих порог 0,01 мм/3ч (а) и 1 мм/3ч (б), в зависимости от заблаговременности для февраля 2014 г. для всего Сочинского региона для моделей COSMO-Ru1 (красная и синяя линии) и COSMO-Ru2 (зеленая и оранжевая линии) для двух методов: ближайшая к станции модельная точка (зеленая и синяя линии) и усреднение модельных данных в радиусе 15 км вокруг станции (оранжевая и красная линии). Начальный срок 00 ч ВСВ.

Отметим, что явления осадков, превышающих порог 10 мм/3ч, вообще не наблюдалось за февраль 2014 г. В целом оценки весьма высоки: для наличия осадков до заблаговременности 30 ч критерий Пирси-Обухова больше 0,4. Интересно, что осреднение в радиусе 15 км хотя и дает некоторое преимущество (для COSMO-Ru2 оранжевая линия по сравнению с зеленой, для COSMO-Ru1 красная линия по сравнению с синей), но не такое значительное и не для всех заблаговременностей, как можно было бы ожидать. Оценка модели COSMO-Ru2 по сравнению с COSMO-Ru1 (зеленая линия по сравнению с синей для метода ближайшей точки, и оранжевая по сравнению с красной для осреднения в радиусе 15 км) показывает, что однозначного преимущества модели более высокого разрешения не наблюдается при традиционной оценке. Чтобы адекватно оценить осадки, требуются новые, так называемые пространственные методы верификации (neighborhood verification (соседские методы верификации), объектно-ориентированные методы, и т. д.) [5, 7].

Программа VERSUS позволяет построить так называемые диаграммы качества (performance diagrams), на которых одновременно представлены предупредительность явления, индекс качества (1 – отношение ложных тревог), сдвиг частоты (пунктирные линии), и показатель угроз (сплошные линии). Идеальный прогноз находился бы в верхнем правом углу: чем ближе к этому углу, тем лучше прогноз. На рис. 3 приведены диаграммы качества для модели COSMO-Ru2 (рис. 3а) и COSMO-Ru1 (рис. 3б) для наличия осадков. Линии означают 95%-ные доверительные интервалы. Видно, что доверительные интервалы широкие, поскольку данных немного. Лучше всего прогнозируются осадки для шага 15 ч. Для шага 24 ч модель COSMO-Ru1 лучше COSMO-Ru2 по всем оценкам, для шага 30 ч – хуже. В целом диаграммы подтверждают вывод о близком качестве обеих версий моделей.

Представляет интерес стратификация оценок по высоте. На рис. 4 представлены оценки ME и RMSE для температуры воздуха на уровне 2 м для COSMO-Ru1 и COSMO-Ru2 для высот около 600, 1000, 1500 и 2000 м, начальный срок 00 ч ВСВ. В таблице приведен список станций, использованных для стратификации.

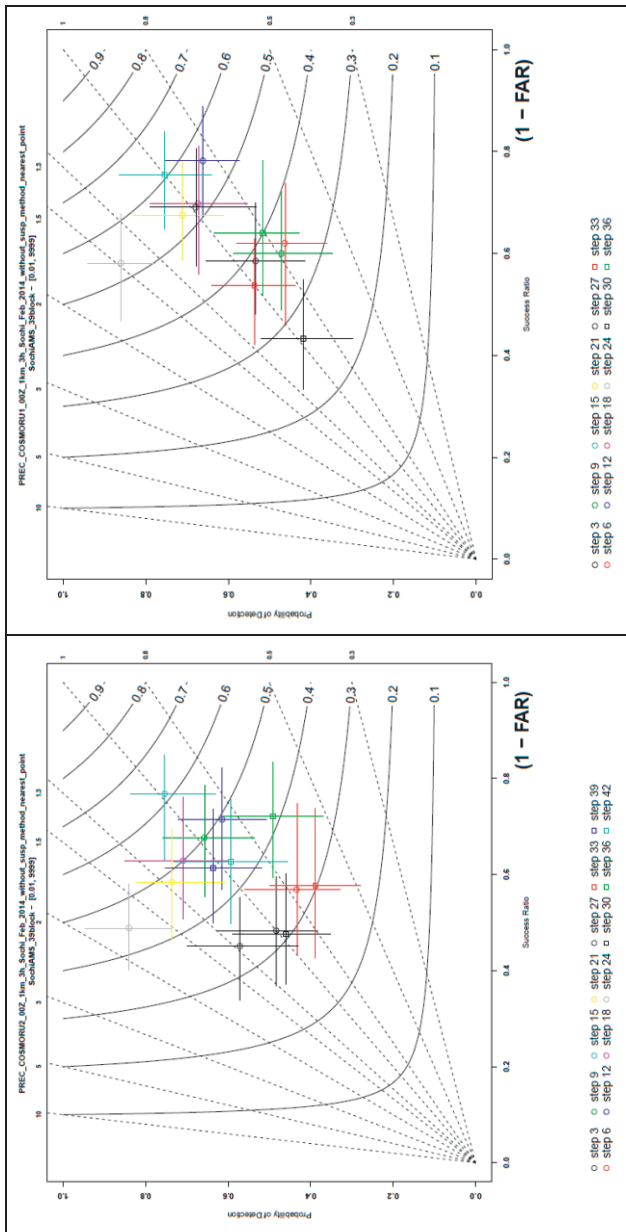
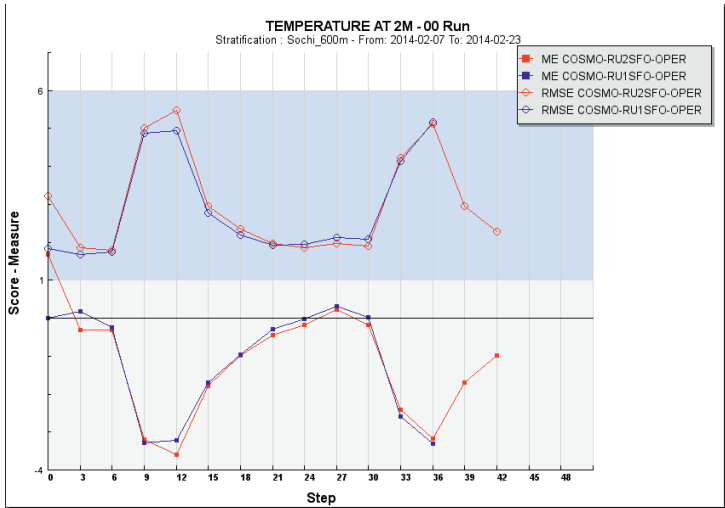
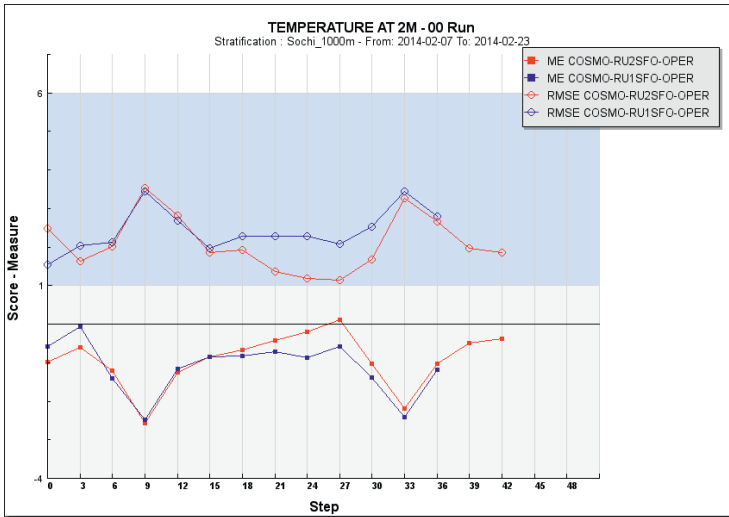


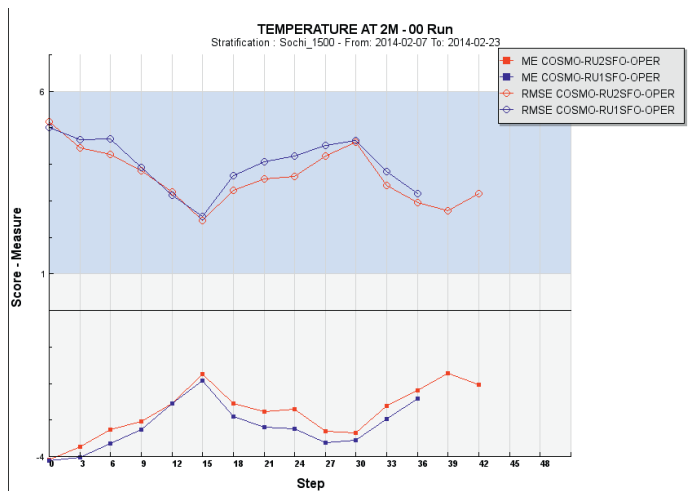
Рис. 3. Диаграммы качества для модели COSMO-Ru2 (а) и COSMO-Ru1 (б) для 3-часовых осадков, превышающих порог 0,01 мм/3ч для февраля 2014 г. для всего Сочиского региона. Линии означают 95%-ные доверительные интервалы. Начальный срок 00 ч ВСВ.



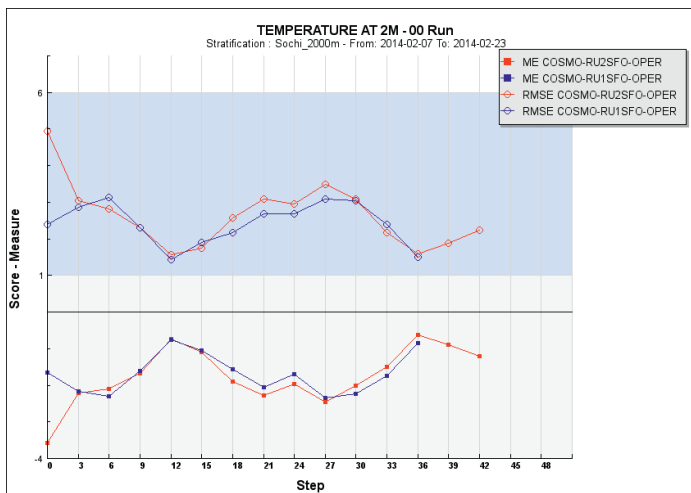
a) 600 m



b) 1000 m



в) 1500 м



в) 2000 м

Рис. 4. Оценки ME и RMSE для температуры воздуха на уровне 2 м для COSMO-Ru1 (синий) и COSMO-Ru2 (красный) для высот около 600 м (а), 1000 м (б), 1500 м (в) и 2000 м (г). Начальный срок 00 ч ВСВ.

**Список станций, использованных для стратификации ошибок
по высоте**

| | |
|--------|---|
| 600 м | 37107-KRASNAYA POLYANA --- LAT:43.683,LON:40.2,HEIGHT:567 39149-Kordon Laura (Roshydromet) --- LAT:43.7,LON:40.265,HEIGHT:570 39159-Sledge-700 ---LAT:43.669,LON:40.289,HEIGHT:701 39557-Nordic-675 ---LAT:43.678,LON:40.237,HEIGHT:676 39558-Nordic-615 ---LAT:43.678,LON:40.234,HEIGHT:615 39560-Ski Jump-650 ---LAT:43.677,LON:40.24,HEIGHT:628 |
| 1000 м | 39153-Snowboard-1025 (IRAM) --- LAT:43.656,LON:40.328,HEIGHT:1027 39154-Freestyle-1080 (IRAM) --- LAT:43.653,LON:40.321,HEIGHT:1077 39165-RKHU-7 (Finish) ---LAT:43.645,LON:40.331,HEIGHT:980 39562-MCarousel-1000 ---LAT:43.668,LON:40.257,HEIGHT:978 |
| 1500 м | 39111-RKHU-4 ---LAT:43.639,LON:40.313,HEIGHT:1580 39136-Alpica-Service-1500(Roshydromet)--- LAT:43.648,LON:40.293,HEIGHT:1465 39161-Biathlon-1500 ---LAT:43.693,LON:40.335,HEIGHT:1497 39162-Biathlon Stadium ---LAT:43.692,LON:40.327,HEIGHT:1471 39163-Ski Stadium ---LAT:43.695,LON:40.329,HEIGHT:1483 39164-Biathlon-1400 ---LAT:43.693,LON:40.319,HEIGHT:1407 39561-MCarousel-1500 ---LAT:43.654,LON:40.251,HEIGHT:1434 |
| 2000 м | 39108-RKHU2 ---LAT:43.63,LON:40.307,HEIGHT:2137 39109-RKHU-3 ---LAT:43.636,LON:40.309,HEIGHT:2043 39146-RKHU-8 ---LAT:43.635,LON:40.308,HEIGHT:1740 |

С увеличением высоты сглаживается суточный ход ошибок (уменьшается дневной максимум ошибок), и в целом имеется тенденция к уменьшению ошибок. Начиная с высоты 1500 м суточный ход меняет знак – в дневное время ошибки становятся меньше, чем в ночное. Для всех высот, кроме 1500 м, характерно уменьшение ошибки в начальный момент времени по модели, связанной с несовершенством начальных полей – для COSMO-Ru1, поскольку эта модель стартует от начальных полей COSMO-Ru2, в которых уже сглажены ошибки начальных данных. Других значимых

отличий между моделями COSMO-Ru1 и COSMO-Ru2 не наблюдается.

Важным преимуществом программы VERSUS является возможность расчета так называемых условных верификаций, т. е. расчет оценок качества какого-либо метеоэлемента при определенных значениях другого метеоэлемента по данным наблюдений или модели. На рис. 5 представлены оценки прогнозов температуры воздуха на уровне 2 м по модели COSMO-Ru1 для станций, расположенных ниже 1200 м, при ясном небе (балл общей облачности <25 %) и облачной погоде (балл общей облачности >75 %). Граница высоты 1200 м была выбрана потому, что выше нее в основном имеется снежный покров, и в модели, и в реальности, и занижение температуры меньше. Из рис. 5 видно, что при облачной погоде качество воспроизведения приземной температуры в дневные часы значительно лучше, чем в ясную погоду. Это говорит о несовершенстве радиационного блока модели, а также процессов на поверхности почвы.

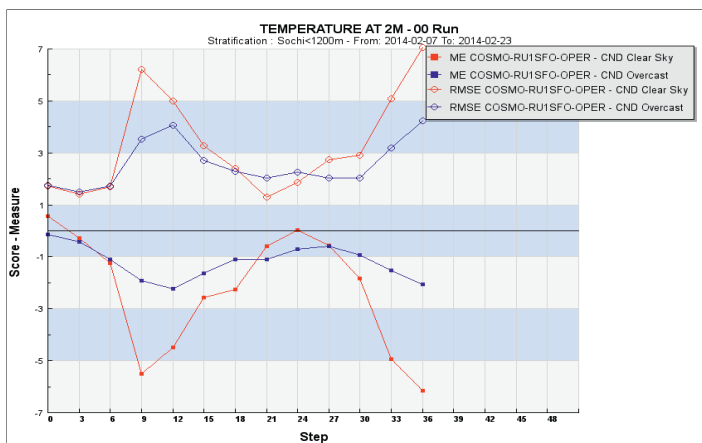


Рис. 5. Оценки прогнозов температуры воздуха на уровне 2 м по модели COSMO-Ru1 для станций, расположенных ниже 1200 м, при ясном небе (балл общей облачности <25 %, красные линии) и облачной погоде (балл общей облачности >75%, синие линии). Начальный срок 00 ч ВСВ.

Исследовательские аспекты верификации, анализ прошлых сезонов – EDI для редких явлений

В [1, 2] проведен подробный анализ качества воспроизведения основных метеозаписей за два прошедших тестовых периода для прибрежного и горного кластеров Олимпийских игр Сочи-2014. Было показано, что средние ошибки могут сильно отличаться от сезона к сезону. Так, например, зимой 2011/2012 гг. модель COSMO-Ru2 в целом превышала приземную температуру воздуха, а теплой зимой 2012/2013 гг. – несколько занижала. Имеющихся тестовых сезонов недостаточно, чтобы дать синоптикам однозначные рекомендации относительно калибровки модельных прогнозов. По этой же причине не удалось сформулировать рекомендации по коррекции прогнозов при различных погодных типах до начала Игр. В то же время анализ прошедших сезонов имеет большое значение для развития модели и для развития методологии верификации в целом. Ниже показано применение некоторых новых индексов оценки качества с акцентом на редких явлениях на примере зимы 2012/2013 гг. Для каждого станционного значения бралась ближайшая модельная точка, сопоставление «модель-станция» делалось с помощью программного комплекса MET [8]. Оценки и графики были построены с помощью программы R.

После 2008 года было предложено несколько новых индексов для оценки редких явлений (extreme dependency score (EDS), stable extreme dependency score (SEDS), extremal dependency index (EDI) and symmetric extremal dependency index (SEDI)), свойства которых подробно обсуждаются в [6], где показано, что EDS и SEDS все же чувствительны к частоте явления, причем EDS не включает доли ложных тревог и, таким образом, может быть улучшен за счет более частого прогноза явления.

В последнем документе ВМО по оценке прогнозов высокого разрешения [9] говорится, что для оценки редких явлений рекомендуется индекс EDI, который при этом дает надежную сравнительную оценку качества прогнозов во всем диапазоне пороговых значений явлений.

Кроме того, в [9] подчеркивается необходимость тщательного исследования данных перед расчетом оценок. Удобный способ для

этого – построение графиков разброса и квантильных графиков, а также гистограмм распределения наблюдений и прогнозов. Такие графики были построены для каждой станции и для каждой заблаговременности прогноза для основных метеозаэментов по Сочинскому региону [2]. С их помощью можно не только проконтролировать качество наблюдений и отбросить плохие данные, но и сделать множество выводов о взаимном распределении модельных и наблюдаемых величин.

Перейдем теперь к анализу индекса качества EDI для всего горного кластера. На рис. 6 представлен временной ход EDI, критерия Пирси-Обухова, а также частоты наблюдаемого явления для часовых осадков, превышающих различные пороговые значения (число случаев часовых осадков, превышающих некоторый порог, суммировалось в трехчасовых интервалах) для модели COSMO-Ru2 с разрешением 2,2 км и американской модели NMMB с разрешением 1 км (для предыдущих тестовых периодов имелись в наличии данные только этой модели помимо модели COSMO). По оси абсцисс отложены пороговые значения: чем больше пороговое значение, тем реже явление. Видно, что по мере приближения частоты явления к нулю критерий Пирси-Обухова также стремится к нулю. Индекс EDI выше, чем критерий Пирси-Обухова, и стремится к нулю медленней, обеспечивая более удобное сравнение моделей. Из рис. 6 можно сделать следующий основной вывод: зимой 2012/2013 гг. COSMO-Ru2 лучше, чем NMMB, прогнозировала факт осадков для горного кластера Сочинского региона, в то же время NMMB лучше предсказывала интенсивные осадки. По подобным графикам легко анализировать частоту явлений, так, например, явление «осадки выше 1 мм/ч» составляет 10 % событий, «выше 2 мм/ч» – 5 %.

Такие же графики были построены для температуры воздуха на уровне 2 м и скорости ветра. Основные выводы следующие. Для температуры по критерию Пирси-Обухова лучше модель COSMO-Ru2, по EDI COSMO-Ru2 лучше вплоть до порога «температура ниже -9°C », затем качество COSMO-Ru2 и NMMB становится сопоставимым. Явление «температура ниже -7°C » наблюдалось в 10 % случаев, а «температура ниже -9°C » – в 5 % случаев.

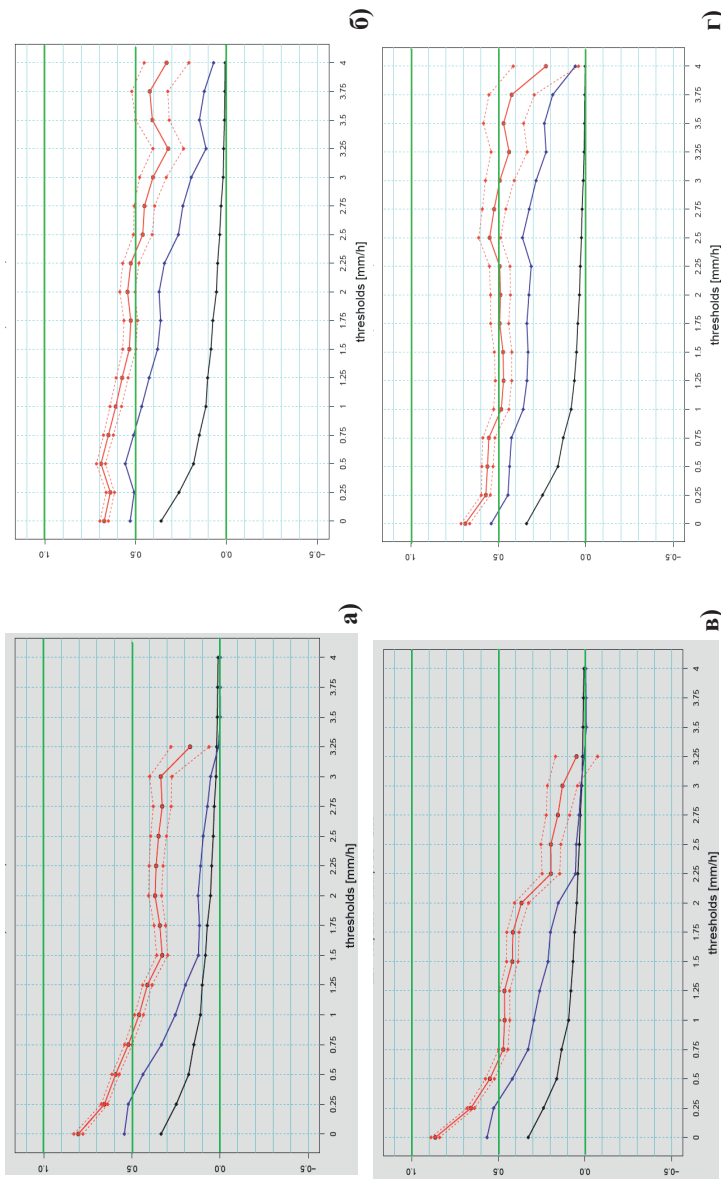


Рис. 6. Временной ход EDI (красная линия), критерия Пирси-Обухова (синяя линия), а также частоты наблюдаемого явления для часовых осадков (черная линия), превышающих различные пороговые значения (по оси абсциссы), для COSMO-Ru2 (слева, затененные графики, а, в) и NMMB (справа б, г).

Для скорости ветра прогнозы оставляют желать лучшего в обеих моделях, хотя оценки для COSMO-Ru2 в целом несколько выше.

Выводы

1. В условиях положительной аномалии температуры и солнечной погоды за первую неделю Олимпийских игр Сочи-2014 все модели занижали приземную температуру воздуха. Это происходит из-за недостатков параметризаций радиационного баланса и подстилающей поверхности. Оценки температуры при условии ясного и облачного неба показывают, что при ясном небе качество моделирования температуры хуже.

2. Модель COSMO-Ru1 с разрешением 1 км оказалась очень полезной для синоптиков за счет возможности получить принципиально новые продукты, как, например, карты линий тока с наложенным полем относительной влажности высокого разрешения. В то же время традиционные точечные оценки качества для этих полей не показывают принципиального улучшения качества COSMO-Ru1 по сравнению с COSMO-Ru2, хотя в целом оценки для COSMO-Ru1 выше, в частности, выше RMSE и ME для приземной температуры воздуха и ME для скорости ветра. Верификацию мезомасштабных моделей высокого разрешения следует дополнять пространственными методами, что станет возможным после завершения работ по созданию совместного анализа стационарных и радарных данных.

3. Критерий Пирси-Обухова для моделей COSMO-Ru1 и COSMO-Ru2 достаточно высок для февраля 2014 г.: для наличия осадков до заблаговременности 30 ч критерий Пирси-Обухова больше 0,4. Осреднение модельных точек в радиусе 15 км вокруг станции дает некоторое преимущество по сравнению с выбором ближайшей к станции модельной точки, но не такое значительное, и не для всех заблаговременностей, как можно было бы ожидать. Оценка модели COSMO-Ru2 по сравнению с COSMO-Ru1 показывает, что однозначного преимущества модели более высокого разрешения не наблюдается при традиционной оценке. Чтобы

адекватно оценить осадки, также требуются новые пространственные методы.

4. Стратификация ошибок температуры воздуха по высоте (600, 1000, 1500 и 2000 м) показала, что с увеличением высоты сглаживается суточный ход ошибок (уменьшается дневной максимум ошибок) и в целом имеется тенденция к уменьшению ошибок. Начиная с высоты 1500 м суточный ход меняет знак – в дневное время ошибки становятся меньше, чем в ночное.

5. Расчет индекса EDI позволяет сравнить качество прогнозов метеоэлементов для высоких пороговых значений (редких явлений), где большинство других индексов близки к нулю. На основе анализа EDI для прогнозов за зиму 2012–2013 гг. для горного кластера Сочи были сделаны следующие выводы: COSMO-Ru2 лучше, чем NMMB, прогнозировала факт осадков, в то время как NMMB лучше предсказывала интенсивные осадки. Для температуры воздуха по критерию Пирси-Обухова лучше модель COSMO-Ru2, по EDI COSMO-Ru2 лучше вплоть до порога «температура ниже $-9\text{ }^{\circ}\text{C}$ », затем качество COSMO-Ru2 и NMMB становится сопоставимым. Для скорости ветра прогнозы оставляют желать лучшего в обеих моделях, хотя оценки для COSMO-Ru2 в целом несколько выше.

Список использованных источников

1. *Муравьев А.В., Бундель А.Ю., Киктев Д.Б., Блинов Д.В., Смирнов А.В.* Верификация мезомасштабных прогнозов в районе Олимпийских игр 2014 г. по первому тестовому периоду. Часть I. Технология верификации и полигонные оценки качества прогнозов на основе модели COSMO // Метеорология и гидрология. – 2013. – № 11. – С. 5–21.

2. *Муравьев А.В., Бундель А.Ю., Киктев Д.Б., Блинов Д.В., Смирнов А.В.* Верификация мезомасштабных прогнозов в районе Олимпийских игр 2014 г. по первому тестовому периоду. Часть II. Диагностические оценки качества и калибровка прогнозов по модели COSMO // Метеорология и гидрология. – 2013. – № 12 – С. 5–22.

3. *Вильфанд Р.М., Ривин Г.С., Розинкина И.А., Астахова Е.Д., Блинов Д.В., Бундель А.Ю., Перов В.Л., Суркова Г.В., Алферов Д.Ю., Казакова Е.В., Кирсанов А.А., Ревокатова А.П., Шатунова М.В., Чумаков М.М.* Негидростатическая система Гидрометцентра России мезомасштабного

краткосрочного прогноза погоды COSMO-Ru / Турбулентность, динамика атмосферы и климата: труды Международной конференции, 13–16 мая 2013 г.– М.: Институт физики атмосферы РАН, 2014.

4. *Шатунова М.В., Ривин Г.С.* Модель с высоким пространственным разрешением COSMO-Ru1: влияние внешних параметров на результаты моделирования // Труды Гидрометцентра России. – 2014. – Вып. 352. – см. настоящий выпуск.

5. *Ebert E.* Fuzzy verification of high resolution gridded forecasts: a review and proposed framework // Meteor. Appl. – 2008. – Vol. 15. – P. 51–64. – doi:10.1002/met.25.

6. *Ferro C.A.T., Stephenson D.B.* Extremal Dependence Indices: improved verification measures for deterministic forecasts of rare binary events // Wea. Forecasting. – 2011. – Vol. 26. – P. 699–713.

7. *Gilleland E., Ahijevych D.A., Brown B.G., Ebert E.E.* Verifying Forecasts Spatially // Bull. Amer. Meteor. Soc. – 2010. – Vol. 91. – P. 1365–1373. – doi: <http://dx.doi.org/10.1175/2010BAMS2819.1>.

8. MET: Model Evaluation Tools, Version 4.0 (METv4.0). User's Guide 4.0.1. / Developmental Testbed Center, 2012, Boulder, Colorado, USA (www.dtcenter.org/met/users)

9. Suggested methods for the verification of precipitation forecasts against high resolution limited area observations / JWGFVR Second Draft, November 2013.

Поступила в редакцию 09.10.2011 г.