

УДК 551.509.321:528.854

**О.В.Ушкаленко, А.Ю.Гаевский**, докт. физ.-мат. наук (Национальный технический университет Украины "КПИ", Киев)

## Статистическое прогнозирование солнечной радиации на основе спутниковых снимков. Алгоритм "вытеснения"

*С целью статистического краткосрочного прогнозирования солнечной радиации по спутниковым снимкам разработана методика, в которой на этапе реконструкции траектории движения облачных элементов используется не традиционный лагранжевый формализм semi-Lagrangian (SL), а предложенный в статье алгоритм "вытеснения", который позволяет существенно сократить время вычислений и не уступает по точности алгоритмам, относящимся к методу SL. Выполнены сравнение и оптимизация различных типов фильтров, предназначенных для сглаживания векторного поля скоростей, а также оптимизация размеров блоков, необходимых для корректного учета корреляции исходных снимков.*

**Ключевые слова:** солнечная радиация, статистическое прогнозирование, векторное поле скоростей, медианный фильтр, кросс-корреляционные функции.

*З метою статистичного короткострокового прогнозування сонячної радіації за супутниковими знімками розроблено методику, в якій на етапі реконструкції траєкторії руху хмарних елементів використовується не лагранжевий формалізм semi-Lagrangian (SL), а запропонований у статті алгоритм "витіснення", який дозволяє суттєво скоротити час розрахунків і не поступається за точністю алгоритмам, що відносяться до методу SL. Виконано порівняння та оптимізацію різних типів фільтрів, призначених для згладжування векторного поля швидкостей, а також оптимізацію розмірів блоків, необхідних для коректного врахування кореляції вихідних знімків.*

**Ключові слова:** сонячна радіація, статистичне прогнозування, векторне поле швидкостей, медіанний фільтр, крос-кореляційні функції.

**Введение и постановка задачи.** Для эффективного использования электроэнергии, получаемой от фотоэлектрических станций (ФЭС), необходимо скоординированное управление генерацией, аккумулярованием и потреблением энергии. Одним из неперемных условий такой координации является прогнозирование ожидаемой мощности станций, которое должно осуществляться на различных временных горизонтах от нескольких минут до двух-трех суток. Чем лучше решаются задачи прогнозирования, тем меньше затраты на поддержание резервных мощностей для восполнения провалов в выработке электроэнергии. Необходимость прогнозирования выработки ФЭС закреплена европейскими стандартами (кодами) в области возобновляемой энергетики, а ошибки в прогнозировании выдаваемой энергии могут караться штрафами [1].

В основе алгоритмов прогнозирования мощности ФЭС лежит расчет прогнозируемой солнечной радиации и учет технических параметров оборудования и расположения компонент

станций. Тип прогноза в значительной степени ориентирован на систему управления и диспетчеризации. Различают прогнозирование на горизонтах от минут до 1 часа (сверхкраткосрочные прогнозы) и от 1 часа до нескольких дней, обычно двух-трех (краткосрочные прогнозы). Для прогнозирования применяются в основном два подхода: первый базируется на физических, а второй – на статистических моделях.

*Физические методы*, за которыми закрепился термин NWP (Numerical Weather Prediction – численное прогнозирование погоды) [2], основаны на решении дифференциальных уравнений гидродинамики, описывающих атмосферные явления и отражающих основные законы физики: сохранение массы, момента, энергии. Переменными являются экспериментально измеряемые величины: скорость ветра, температура, давление, влажность, поток солнечного излучения и др. Атмосферная область представляется в виде слоев, имеющих одинаковую латеральную координатную сетку. Решения вычисляются послойно

с учетом граничных условий между слоями и на границах рассматриваемой атмосферной области. Реализация NWP-расчетов требует значительных компьютерных ресурсов, объем которых определяется в первую очередь размером ячеек координатной сетки и числом слоев (их может быть несколько десятков). Физический подход предпочтителен на горизонтах от нескольких часов до нескольких дней, поскольку он включает пространственные и временные факторы в рамках общей гидродинамической модели. Несмотря на возможность дальнего прогнозирования, физическому подходу свойственны некоторые ограничения, а именно: недостаточность экспериментальных данных для калибровки и параметризации, а также не очень высокое разрешение при большой площади рассчитываемого домена и трудности с учетом топографии.

В отличие от физического подхода, в *статистических методах* множество энергетических зависимостей и метеорологических процессов точно не анализируются, однако устанавливается связь между "историческими" данными о вырабатываемой энергии и погодой. На основе такой статистической связи составляется прогноз о предстоящей выработке, и входные данные могут быть преобразованы в выходную мощность при использовании незначительных вычислительных ресурсов. К статистическим относятся [3] авторегрессионная модель скользящего среднего (ARMA – Autoregressive Moving Average), авторегрессионная интегрированная модель скользящего среднего (ARIMA – Autoregressive Integrated Moving Average), а также методы, основанные на фильтрах Кальмана (Kalman Filters), искусственных нейронных сетях (ANN – Artificial Neural Networks), гауссовых процессах (GP – Gaussian Process) и др. Поскольку ветер быстро меняется во времени, статистические модели, как правило, эффективны для очень краткосрочных прогнозов (1-3 часа). На таких небольших временных горизонтах точность статистических расчетов радиации может даже превышать точность NWP-расчетов.

Отдельным направлением в статистическом прогнозировании являются методы, основанные на использовании *спутниковых данных* в виде снимков земной поверхности и облачного покро-

ва, получаемые в видимом и инфракрасном диапазонах [4–8]. Для реализации этих методов не требуются обширные архивные данные (как в ARMA-расчетах или в расчетах, базирующихся на экспертных системах), поскольку анализ основывается на текущих тенденциях движения облачных масс. Исходными данными является серия последовательных снимков (обычно двух) облачного покрытия, выполненных с небольшим интервалом (15-30 минут), а также снимок той же территории без облаков, отображающий земное альbedo. Эти данные обрабатываются с помощью какого-либо из статистических методов распознавания изображений с последующей фильтрацией, и в результате получается векторное поле скоростей (MVF – Motion Vector Field), отвечающее облачной адвекции [7, 8]. На основе рассчитанного векторного поля генерируется распределение индекса облачности через определенные временные интервалы прогнозирования. Осуществляется это путем расчета траекторий элементов облаков с помощью решения дифференциальных уравнений движения в координатах Лагранжа (схема semi-Lagrangian [9]). Существуют две схемы решения этих уравнений, отличающиеся направлением выполнения расчета, это схемы "up-stream" и "down-stream". В первой схеме [9] для каждого текущего элемента прогнозируемого изображения, относящегося к моменту времени  $t$ , определяется элемент на предшествующем изображении для момента  $t - \Delta t$ , который мог попасть в точку текущего элемента. Делается это путем вычисления траекторий, обращенных назад во времени, поэтому эту схему называют "up-stream" – движение вверх по потоку. В схеме расчета "down-stream", называемой также "forwardsemi-Lagrangian" (FSL) [10], исходят из последнего экспериментально полученного изображения в момент  $t - \Delta t$  и рассчитывают траектории элементов в координатах Лагранжа на основе имеющегося поля скоростей. При этом приходится учитывать деформацию элементов ввиду неоднородности MVF. Следует отметить, что обе схемы расчета в рамках метода SL требуют значительного машинного времени.

В данной работе предложен иной, более простой алгоритм расчета прогнозируемого изобра-

жения, реализуемый в кинематической схеме Эйлера, т.е. без учета деформации координатных сеток [11], и названный нами алгоритмом "вытеснения". Этот алгоритм является быстроедействующим и, как следует из результатов данной работы, расчеты прогноза, выполненные с его помощью, не уступают в точности расчетам методом SL. Метод "вытеснения" позволяет получать прогнозируемые изображения и соответственно значения солнечной радиации в заданной географической точке на временных горизонтах до 75-90 минут с погрешностью, не превышающей 13-15% для относительно однородных облаков.

**Основные формулы и алгоритмы вычислений.** При расчетах статистическим методом в качестве входных данных используются изображения видимого диапазона, полученные с геостационарного спутника METEOSAT-8 (рис. 1). Информация об облачном покрове и альбедо получается путем обработки спутниковых изображений методом Heliosat [4], в результате расчет дает карты распределения индекса облачности.

На основе двух последовательных (во времени) карт индекса облачности рассчитывается векторное поле скоростей (MVF) [5–8]. Следует отметить, что в большинстве алгоритмов расчета MVF пиксельное изображение разбивается на квадратные блоки достаточно большой величины (например, 13×13 пикселей), чтобы их можно было идентифицировать при сравнении двух

последовательных изображений. Для расчета поля скоростей нами в работе применялись два метода распознавания изображений: это метод отображения блоков (BM – Block Matching) [12] и метод кросс-корреляционных функций (CCF – Cross-Correlation Functions) [8]. Название первого указывает на способ сравнения двух изображений: последовательно просматривается каждый блок первого изображения с распределением интенсивности  $I^{(1)}(r)$  (где  $r$  – координата блока в изображении), и во втором изображении ищется блок  $I^{(2)}(r+d)$  наиболее похожий на блок  $I^{(1)}(r)$  первого изображения. Из текущего блока в найденный блок проводится вектор смещения  $d$ , который и является одним из векторов MVF. Критерием схожести блоков может быть, например, минимальное значение средней абсолютной разности (MAD – Mean Absolute Difference), которая вычисляется как

$$MAD \equiv \left\langle \left| I^{(1)}(r) - I^{(2)}(r+d) \right| \right\rangle = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \left| I_{ij}^{(1)} - I_{ij}^{(2)} \right|, \quad (1)$$

где  $I_{ij}^{(1)}$ ,  $I_{ij}^{(2)}$  – интенсивности пикселя  $ij$  в первом и во втором изображениях соответственно; суммирование в (1) выполняется по всем  $N \times N$  пикселям каждого блока. Существуют различные алгоритмы оптимизации процесса поиска похожих блоков, которые составляют содержание технологии, называемой Optical Flow [13].

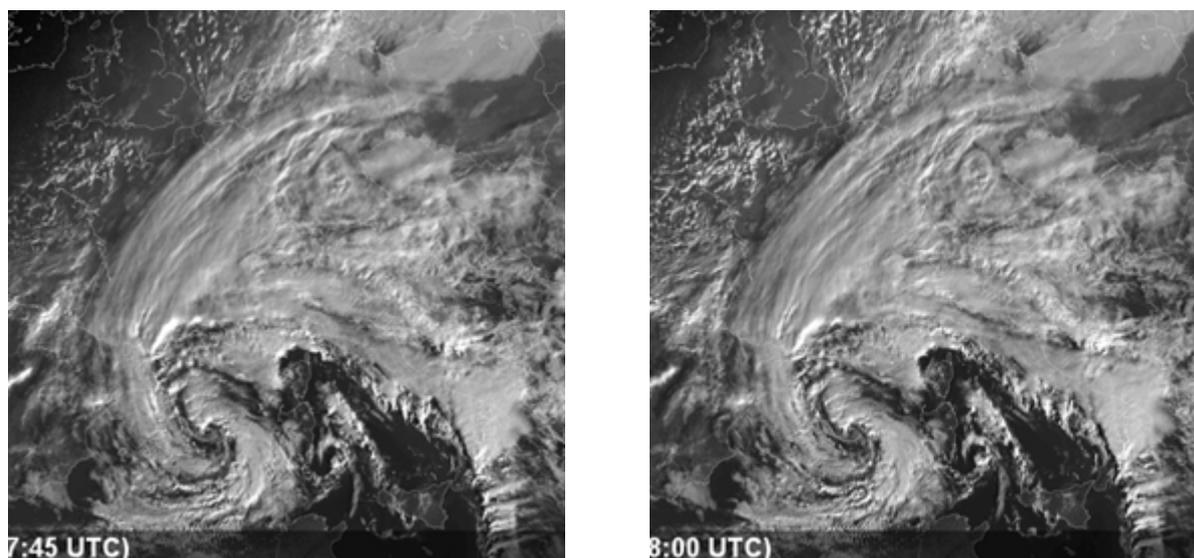


Рис. 1. Пример двух исходных последовательных изображений (19 ноября 2013 года 08:45UTC и 09:00UTC).

Согласно методу ССФ [8] выбор двух похожих блоков осуществляется путем сравнения корреляторов второго порядка вида:

$$CCF \equiv \langle I^{(1)}(r)I^{(2)}(r+d) \rangle = \frac{\sum_{i=0}^N \sum_{j=0}^N (I_{ij}^{(1)} - \overline{I^{(1)}})(I_{ij}^{(2)} - \overline{I^{(2)}})}{\sqrt{\sum_{i=0}^N \sum_{j=0}^N (I_{ij}^{(1)} - \overline{I^{(1)}})^2} \cdot \sqrt{\sum_{i=0}^N \sum_{j=0}^N (I_{ij}^{(2)} - \overline{I^{(2)}})^2}} \quad (2)$$

и поиска вектора смещения  $d$ , обеспечивающего максимальное значение ССФ. В формуле (2)  $\overline{I^{(1)}}$ ,  $\overline{I^{(2)}}$  – средние значения интенсивностей для текущего блока первого и второго изображений соответственно. Как показали расчеты, результаты определения полей смещений обоими методами (ВМ и ССФ) отличались незначительно.

При расчетах поля скоростей обычно используются два допущения: во-первых, считается, что интенсивности пикселей остаются постоянными при перемещении, во-вторых, соседние векторы скоростей не могут значительно различаться по модулю и по направлению. Однако получаемое векторное поле может содержать "выбросы" – векторы, сильно отличающиеся от своих соседей, что не согласуется со вторым допущением (рис. 2а), и поэтому требуется сглаживание поля с помощью какого-либо типа фильтра.

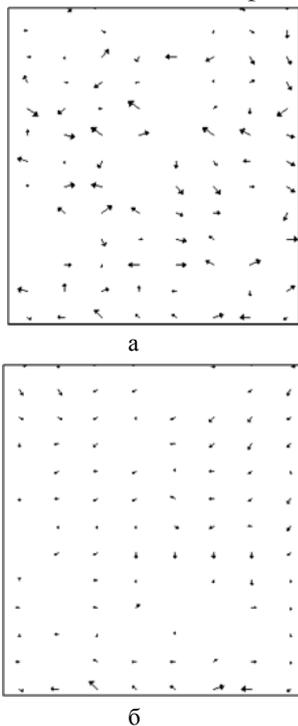


Рис. 2. Пример векторного поля скоростей до (а) и после (б) сглаживания.

Оптимизация по формулам (1) либо (2) с последующей фильтрацией дает лишь векторы скоростей в узлах крупномасштабной сетки блоков. Для получения векторов для каждого пикселя необходима дополнительная процедура, которая связана с определением радиуса влияния блоков. Для построения поля векторов скоростей для каждого пикселя нами был использован объектный метод Крессмана [14].

Прогнозируемое изображение получается путем расчета траекторий движения элементов изображения. В настоящей работе были применены два алгоритма получения такого изображения: "up-stream" SL [9] и предложенный в настоящей работе алгоритм "вытеснения". В основе первого алгоритма лежит уравнение непрерывности потока интенсивности  $I(r, t)$ :

$$\frac{dI}{dt} = \frac{\partial I}{\partial t} + \frac{dr}{dt} \nabla I = 0, \quad (3)$$

$$\text{где } \frac{dr}{dt} = V(r) \quad (4)$$

– вектор скорости на карте MVF. Уравнение (3) отвечает сохранению скалярной величины  $I$  вдоль траектории движения частицы (в нашем случае – элемента облака). Обозначим узлы сетки, для которых определены векторы MVF, как  $m=(ij)$ . Пусть на прогнозируемом изображении, относящемся к моменту времени  $t$ , частица в этих узлах (точках "прибытия") описывается функцией  $I(r_m, t)$ . Тогда на предшествующем исходном изображении в момент  $(t - \Delta t)$  эта частица находилась в положении  $(r_m - d_m)$ , где  $d_m$  – вектор смещения. Соответственно на первом исходном изображении (момент  $t - 2\Delta t$ ) данная частица имела координату  $(r_m - 2d_m)$  (точка "отправления"). Уравнение сохранения функции  $I(r_m, t)$  (1) в дискретном представлении выглядит так:

$$\frac{I(r_m, t) - I(r_m - 2d_m, t - 2\Delta t)}{\Delta t} = 0. \quad (5)$$

Смещение  $d_m$  ищется путем решения следующего дискретного уравнения, которое с точностью до членов  $o(\Delta t^2)$  является следствием уравнения (4):

$$d_m = \Delta t V(r_m - d_m, t - \Delta t). \quad (6)$$

Итерационная процедура выглядит как

$$d_m^{(k+1)} = \Delta t V(r_m - d_m^{(k)}, t - \Delta t), \quad (7)$$

где  $(k+1)$  и  $k$  обозначают текущую и предыдущую итерации. Вычисление интенсивности  $I(r_m, t)$  в точках "прибытия" в схеме "up-stream" начинается с решения уравнений (7) путем итераций для каждого узла  $m$ , в результате находят-ся смещения  $d_m$ . При этом компоненты двумер-ных векторов скорости вычисляются интерполя-ционным образом через компоненты векторов в соседних узлах. Исходя из найденных смещений  $d_m$ , определяются также путем интерполяции значения  $I(r_m - 2d_m, t - 2\Delta t)$  в точках "отправле-ния" и вычисляются интенсивности  $I(r_m, t)$  из уравнения (5).

Второй алгоритм нахождения прогнозируе-мого изображения, примененный в данной рабо-те, состоит в применении к последнему экспери-ментально полученному изображению поля век-торов смещений  $d_m$  пикселей, найденных мето-дом Крессмана согласно векторам смещений блоков. При этом расчет выполняется в коорди-натах Эйлера, т.е. положение всех частиц привя-зано к двумерной сетке пикселей. Суть алгоритма "вытеснения" состоит в двух правилах:

- если в результате применения карты MVF к исходному изображению две или более частиц попадают в один узел нового прогнози-руемого изображения, то в этот узел помещается та частица, которая имеет наибольшую яркость;
- если же какой-либо из узлов оказывается не занятым новой частицей, то сохраняется прежний пиксель в этом узле, т.е. интенсивность узла не меняется.

Точность прогнозирования оценивалась пу-тем вычисления отклонений прогнозного изобра-жения от фактически наблюдавшегося изображе-ния. Оценка выполнялась по трем параметрам:

- среднеквадратичной ошибке RMSE:

$$RMSE = \sqrt{\frac{1}{M} \sum_{m=1}^M (I_{pred,m} - I_{fact,m})^2}; \quad (8)$$

- средней ошибке смещения MBE:

$$MBE = \frac{1}{M} \sum_{m=1}^M (I_{pred,m} - I_{fact,m}); \quad (9)$$

- средней абсолютной ошибке MAE:

$$MAE = \frac{1}{M} \sum_{m=1}^M |I_{pred,m} - I_{fact,m}|, \quad (10)$$

где индекс  $m$  под знаками сумм нумерует пары значений индекса облачности, которые относятся к  $M$  блокам изображений в прогнозируемый момент времени  $t$ . Величины  $I_{pred,m}$  и  $I_{fact,m}$  в (8)–(10) – это предсказанное и фактическое (т.е. вычисленное по спутниковому снимку в момент  $t$ ) значения ин-декса облачности. Блоки  $m$  выбирались случай-ным образом, а значения  $I_{pred,m}$ ,  $I_{fact,m}$  определя-лись путем усреднения внутри каждого блока. Очевидно,  $M$  было меньше полного числа бло-ков, на которые разбивалось изображение.

**Результаты расчетов.** Предварительными этапами для расчета прогноза радиации являются следующие: определение размера блоков, между которыми вычисляется корреляция; оптимизация характеристик сглаживающих фильтров; задание радиуса влияния Крессмана при построении MVF. При настройке параметров фильтров суще-ственным оказался выбор области действия фильтра (размера матрицы свертки). Например, для медианного фильтра минимальное значение ошибки RMSE (8) наблюдалось при размере мат-рицы свертки, равном 15 (рис. 3).

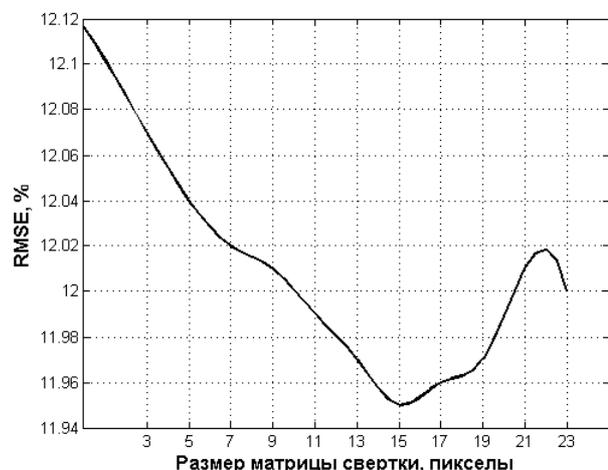


Рис. 3. Зависимость RMSE от размера матрицы свертки для медианного фильтра.

Для матрицы свертки  $15 \times 15$  были выполнены расчеты прогнозируемых снимков при использовании трех типов фильтров: гауссова, медианного и фильтра по среднему значению. Результаты расчетов отличались по критерию RMSE незначительно (менее 0,5%); но более предпочтительным на горизонтах прогноза  $\geq 1$  часа оказался медианный фильтр (рис. 4).

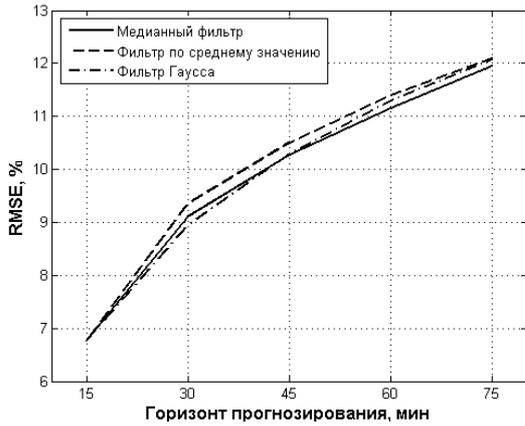


Рис. 4. Ошибка RMSE (8) в зависимости от горизонта прогнозирования для фильтров по среднему значению, медианного и гауссова (матрица свертки  $15 \times 15$ ).

Следует отметить существенную зависимость ошибок прогнозирования от размера блоков, на которые разбивается изображение и по которым ведется расчет корреляторов (2). Как и ожидалось, ошибки велики при малом размере блоков, но они уменьшаются с увеличением блоков, достигая некоторого нижнего предельного значения (рис. 5). Как следует из графиков на рисунке 5, предложенный нами метод прогнозирования является более предпочтительным, чем SL, особенно при небольших размерах блоков. Корректный учет корреляции в методе "вытеснения" возможен уже при размере блоков  $9 \times 9$  или  $11 \times 11$  пикселей.

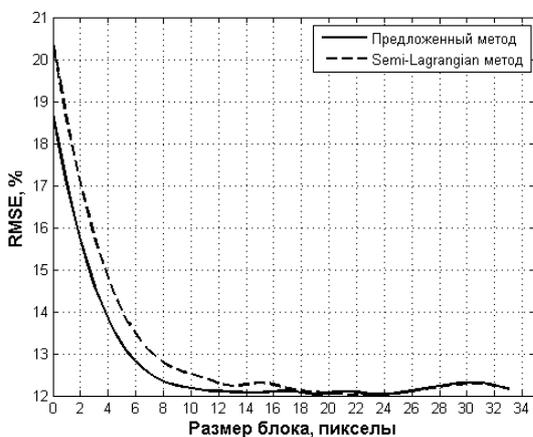
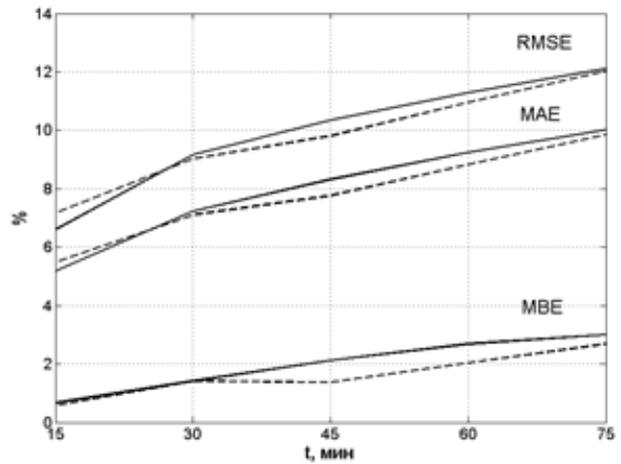
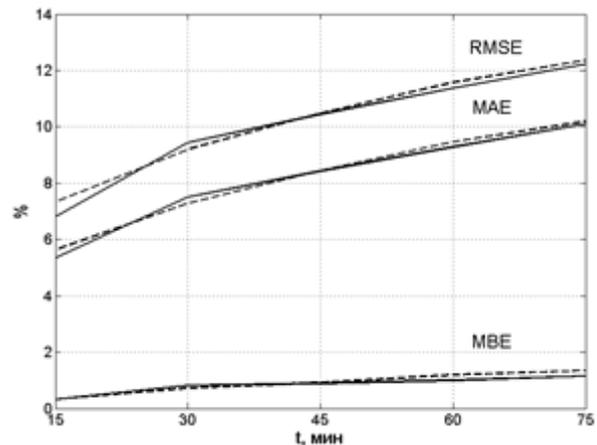


Рис. 5. Зависимость RMSE от размера блоков в исходных изображениях при расчетах двумя методами: SL и предложенным методом.

Итоговое сравнение двух методов прогнозирования приведено на рис. 6, где показаны зависимости ошибок (8)–(10) от горизонта прогноза. Различие в качестве прогнозов между предлагаемым методом и методом SL не является существенным и не превышает статистической погрешности самих расчетов. Свидетельство этому – сближение результатов расчетов при увеличении объема выборки: кривые на рис. 6б при  $M=1600$  почти сливаются. Однако метод "вытеснения" более предпочтителен, чем SL, ввиду меньших требований к вычислительным ресурсам.



а



б

Рис. 6. Зависимость ошибок (8)–(10) от горизонта прогнозирования для предложенной схемы расчета траекторий (сплошные линии) с медианным фильтром и для схемы SL (штриховые линии): а – число учитываемых блоков при подсчете ошибки  $M=1000$ ; б –  $M=1600$ .

**Выводы.** Оптимальным при расчете полей скоростей (карт MVF) для краткосрочного прогнозирования солнечной радиации по спутниковым снимкам на горизонтах 1-1,5 часа является

применение медианного фильтра с матрицей свертки порядка  $15 \times 15$ .

Предложенный алгоритм "вытеснения" для расчета прогнозируемого распределения индекса облачности не уступает в точности традиционным алгоритмам "up-stream" и "down-stream" метода SL, но выигрывает в быстродействии и меньших требованиях к вычислительным ресурсам.

Для корректного учета корреляции при сравнении исходных спутниковых снимков и дальнейшего расчета прогноза методом "вытеснения" достаточен размер блоков  $9 \times 9$  пикселей.

1. <http://pjm.com/~media/committees-groups/task-forces/irtf/postings/pris-task3b-best-practices-from-other-markets-final-report.ashx>.

2. Kalnay E. Atmospheric modeling, data assimilation, and predictability. Cambridge University Press, 2002, 341pp.

3. [http://wire1002.ch/fileadmin/user\\_upload/Restricted/Current\\_State\\_Report/Current\\_State\\_Report\\_draft\\_final.pdf](http://wire1002.ch/fileadmin/user_upload/Restricted/Current_State_Report/Current_State_Report_draft_final.pdf).

4. Hammer A., Heinemann D., Hoyer C., Kuhlemann R., Lorenz E., Müller R. and Beyer H. G. Solar energy assessment using remote sensing technologies // Remote sensing of environment. – 2003. – Vol. 86. – No. 3. – P. 423–432.

5. Bannehr L., Rohn M. and Warnecke G. A functional analytic method to derive displacement vector fields from satellite image sequences // International Journal of Remote Sensing. – 1996. – Vol. 17. – No. 2. – P. 383–392.

6. Cano D., Monget J.M., Albussion M., Guillard H.,

Regas N., Wald L. A method for the determination of global solar radiation from meteorological satellite data // Solar Energy. – 1986. – 37. – No. 1. – P. 31–39.

7. Hammer A., Heinemann D., Lorenz E., Lücke B. Short-term forecasting of solar radiation: a statistical approach using satellite data // Solar Energy. – 1999. – Vol. 67. – No. 1-3. – P. 139–150.

8. Hamill T.M., Neukorn T. A short-term forecast scheme using cross correlations // Weather and Forecasting. – 1993. – Vol. 8. – No. 4. – P. 401–411.

9. Staniforth A., Cote J. Semi-Lagrangian integration schemes for atmospheric models – A review // Monthly Weather Review. – 1991. – Vol. 119. – No. 9. – P. 2206–2223.

10. Nair R.D., Scroggs J.S., Semazzi F.H.M. A forward-trajectory global semi-Lagrangian transport scheme // Journal Comput. Physics. – 2003. – Vol. 190. – No. 1. – P. 275–294.

11. Гаевский А.Ю., Ушкаленко О.В. Эффективная схема расчета двумерной адвекции в задачах прогнозирования солнечной радиации. "Відновлювана енергетика XXI століття". Матеріали XV міжнародної конференції – Київ, 2014, с.178 – 180.

12. Lu J., Liou M.L. A simple and efficient search algorithm for block-matching motion estimation // IEEE Trans. Circuits And Systems For Video Technology. – 1997. – Vol. 7. – No. 2. – P. 429–433.

13. Barjatya A. Block matching algorithms for motion estimation.

<http://www.mathworks.com/matlabcentral/fileexchange/8761-block-matching-algorithms-for-motion-estimation>.

14. Cressman G.P. An operational objective analysis system // Monthly Weath.Rev. – 1959. – Vol. 87. – No. 10. – P. 367–374.