

В.И. Акселевич

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ АВТОРЕГРЕССИОННОЙ МОДЕЛИ ДЛЯ ПРОГНОЗА ТЕМПЕРАТУРЫ

Рассматриваются результаты прогноза температуры на различных уровнях и в зависимости от различных вариантов работы алгоритма. Производится сравнение результатов с инерционным и климатическим прогнозом. Выявляются оптимальные условия работы алгоритма. Обсуждаются варианты исходных данных для прогноза и возможности получения информации о вертикальных профилях температуры при различных ситуациях.

Проблема разработки прогноза вертикального распределения температуры имеет существенное значение, поскольку полученная информация может использоваться для решения ряда прикладных задач. Среди них:

- обеспечение надежной работы и сохранения заданной мощности авиационных и ракетных двигателей;
- исследование свойств и характеристик электромагнитного излучения различной природы;
- определение температурной стратификации и устойчивости атмосферы, от чего зависит успех прогноза погоды и загрязнения окружающей среды;
- климато-экологический мониторинг природной среды;
- дешифрирование спутниковой информации;
- прогноз обледенения, болтанки, конденсационных следов и положительного отклонения температуры воздуха от стандартной;
- расчет лидарных трасс и оптимальных условий обеспечения надежной связи;
- расчет траекторий полета различных тел;
- обеспечение работоспособности и надежности функционирования аэростатов.

Как правило, выделяют три различных методических подхода к прогнозу температуры воздуха. Первый – восстановление из данных о геопотенциале, полученных путем численного прогноза [1]. Второй – применение интегрирования уравнения теплопереноса [2]. Третий – использование методов статистической интерполяции по другим прогностическим полям [3, 4].

На основе этих подходов и с учетом их различных сочетаний для прогноза вертикальных профилей температуры используют ряд методик, различающихся по области применения и по набору необходимой исходной информации.

В [5] рекомендуется использовать траекторную методику прогноза температуры, которая наиболее приспособлена к имеющемуся стандартному набору исходных данных, апробирована Роскомгидрометом и имеет достаточно высокую оправданность. Вместе с тем существует ряд задач, когда подобный

традиционный подход не является правомерным. Например, часто возникает проблема восполнения информации, если имеются пропуски в данных наблюдений и необходимы их интерполяция и экстраполяция. Достаточно остро стоит проблема восстановления и прогноза вертикальных профилей температуры в средней и верхней стратосфере с использованием физико-статистических методов. Нельзя исключать и вариантов потери или невозможности получения метеоинформации по различным причинам. Поэтому наряду с традиционным направлением прогнозирования вертикального распределения температуры развиваются и альтернативные на основе применения метода аналогов, метода группового учета аргументов, методики типизации вертикальных профилей и использования авторегрессионных схем и уравнений.

Рассмотрим подробнее авторегрессионную модель для прогноза значения температуры на заданном реперном уровне. Физически подобный подход оправдан определенной цикличностью изменения состояния атмосферы, наличием волновых движений и законом сохранения энергии.

Целесообразно представить изменение температуры в качестве стационарного процесса и описывать его с помощью известной модели авторегрессии:

$$E\{y[t]\} = \sum_{\tau=1}^{S(0)} Q E\{y[t-\tau]\},$$

где $y[t]$ – характеристики исследуемого процесса в дискретные моменты времени $t = 1, 2, \dots, n$; $y[t] = E\{y[t]\} + \varepsilon[t]$ – случайные ошибки измерений с нулевым математическим ожиданием и конечной дисперсией; Q – параметры модели; $S(0)$ – сложность или порядок модели.

Качество оценивания нужно связывать с прогнозированием. При этом в силу прямой зависимости ошибки прогноза от ошибок оценки параметров следует стремиться к получению оценки Q как можно ближе к истинным значениям Q . Получение оценок параметров авторегрессии целесообразно основывать на минимизации отклонений от гиперплоскости

авторегрессии одновременно вдоль всех направлений, задаваемых $y[t]$ и $y[t-\tau]$. Вместе с оцениванием параметров m необходимо также решать задачу селекции оптимальной структуры авторегрессионной модели порядка S^* . Это сводится к оцениванию параметров $Q(S)$ при различных S и выбору такого S^* , при котором достигается минимум ошибки прогноза.

На практике необходимо провести адаптацию модели к реальным условиям и исследовать чувствительность модели к изменению отдельных параметров.

Модель авторегрессии позволяет варьировать длину исходного ряда наблюдений, количество шагов прогноза, максимальный порядок авторегрессии, критерий оценивания (1 – среднеквадратическая ошибка прогноза на контрольной выборке; 2 – финальная ошибка прогнозирования; 3 – среднеквадратическая ошибка прогноза на перекрестной выборке), количество сохраняемых моделей, номер модели, используемой для прогноза, длину контрольной выборки в случае выбора первого критерия оценивания.

Для оценки работы алгоритма варьировали все эти параметры в различных сочетаниях, добиваясь лучших результатов прогноза на период от 1 до 6 сроков будущих наблюдений. При этом количество исходных элементов изменяли от 10 до 79, максимальный порядок авторегрессии от 1 до 19, число сохраняемых лучших моделей от 1 до 29. Поскольку первый критерий оценивания предполагает заданную величину контрольной выборки, исследовалось изменение и этой величины, также влиявшей на конечный результат прогноза. Кроме того, был проведен эксперимент с выбором номера модели, который изменялся от 1 до 29. В качестве исходных данных использовали сами значения температуры и их отклонения от «нормы», под которой понимали известное из климатических справочников среднее для данного периода времени значение температуры для пункта наблюдения. По итогам расчетов были сделаны следующие выводы:

1. Результат не зависит от числа сохраняемых моделей. Наименьшая ошибка в ходе вычислительного эксперимента соответствовала 4-й модели. Поэтому необходимо сохранять не менее 4 моделей.

2. Максимально возможное количество моделей зависит от порядка авторегрессии. При его значении, равном 1, это число составляет 1, 2 – 3, 3 – 7, 4 – 15 и т.д.

3. Успешность прогноза более зависит от номера модели, чем от порядка авторегрессии. Оптимальным порядком с точки зрения конечного результата оказался 5-й.

4. При изменении количества исходных наблюдений наилучшие результаты варьировались в зависимости от шага прогнозирования. При прогнозе на один–два шага наилучшие результаты получились при наличии 30 элементов в исходной выборке, 3 – 4 шага – 25 ... 30, 5 шагов – 40, а 6 – 20 элементов.

5. Удобнее всего использовать для текущей оценки качества прогнозирования второй критерий – финальную ошибку прогнозирования.

Таким образом, модель оказывается наиболее чувствительной к выбору номера модели и порядка авторегрессии. Несколько меньше зависит успешность прогноза от количества исходных наблюдений и замены абсолютных значений отклонениями от норм.

После выявления оптимальных параметров для периодов 12, 24, 36, 48, 60, 72 ч без учета суточного хода и 1–6 сут с его учетом для дня и ночи определялись значения температуры и отклонения температуры от климатической нормы на уровнях Земли, 700, 500 и 300 гПа на станции Кефлавик для зимы и лета по 10-летней выборке. В табл. 1 представлены результаты прогноза температуры (в виде средней абсолютной ошибки прогноза и среднеквадратической ошибки прогноза) по авторегрессионной модели для июля (по десятилетней выборке) с интервалом 12 ч для станции Кефлавик. Оценить результаты прогноза позволяет табл. 2, где представлены относительные среднеквадратические ошибки прогноза температуры, полученные с помощью сравнения среднеквадратической ошибки прогноза по авторегрессионной модели с климатическим среднеквадратическим отклонением [6] и среднеквадратической ошибкой инерционного прогноза.

Анализ табл. 1 показывает, что наименьшие средние значения абсолютной ошибки прогноза на всех поверхностях, кроме 500 гПа, соответствуют заблаговременности 12 ч, а наибольшие – 72 ч, но при этом на уровне Земли и 500 гПа не наблюдается монотонного возрастания ошибки. Если же сравнить величину ошибок на различных уровнях, то наименьшие их значения соответствуют уровню Земли, а наибольшие, кроме заблаговременности 12 ч, поверхности 300 гПа. Аналогичные результаты характерны и для среднеквадратической ошибки прогноза, за исключением поверхности 500 гПа. Анализ представленной информации позволяет сделать вывод, что для любого срока прогноз с помощью метода авторегрессии дает существенный выигрыш относительно прогноза по климатическому справочнику. Вместе с тем по сравнению с инерционным прогнозом по тем же исходным данным модель авторегрессии имеет преимущество только на уровнях 500 и 300 гПа. При этом наибольший выигрыш получен при заблаговременности прогноза 36–48 ч. Очевидно, модель следует несколько доработать. Использование отклонений принесло худшие результаты, чем работа с самими значениями температуры. Основные надежды на улучшение качества прогноза при использовании авторегрессионной модели связаны с изменением вида исходных данных. В их качестве предполагается целесообразным использовать коэффициенты разложения по естественным ортогональным составляющим (ЕОС), предлагаемые в [7, 8].

Таблица 1

**Результаты прогноза температуры на уровнях Земли, 700, 500 и 300 гПа
(средняя абсолютная ошибка/среднеквадратическая ошибка, °С)**

Уровни	Заблаговременность прогноза, ч					
	12	24	36	48	60	72
Земля	<u>1,4</u>	<u>1,5</u>	<u>1,9</u>	<u>1,8</u>	<u>1,9</u>	<u>1,9</u>
	0,9	1,0	1,1	1,0	1,2	1,3
700 гПа	<u>1,7</u>	<u>2,2</u>	<u>2,4</u>	<u>2,8</u>	<u>3,1</u>	<u>3,4</u>
	1,3	1,8	1,8	2,4	2,4	2,7
500 гПа	<u>3,0</u>	<u>3,1</u>	<u>2,8</u>	<u>2,7</u>	<u>3,2</u>	<u>3,5</u>
	2,0	2,4	1,8	1,9	2,7	3,1
300 гПа	<u>2,4</u>	<u>3,1</u>	<u>3,4</u>	<u>3,4</u>	<u>3,5</u>	<u>3,6</u>
	1,9	1,9	2,0	2,0	2,3	2,4

Таблица 2

**Относительные среднеквадратические ошибки прогноза температуры на уровнях Земли, 700, 500 и 300 гПа
(отношение ошибки модели к климатическому среднеквадратическому отклонению/отношение ошибки модели
к среднеквадратической ошибке инерционного прогноза, %)**

Уровни	Заблаговременность прогноза, ч					
	12	24	36	48	60	72
Земля	<u>39</u>	<u>43</u>	<u>48</u>	<u>43</u>	<u>52</u>	<u>57</u>
	82	111	100	125	120	217
700 гПа	<u>36</u>	<u>50</u>	<u>50</u>	<u>67</u>	<u>67</u>	<u>75</u>
	108	120	95	120	126	142
500 гПа	<u>50</u>	<u>60</u>	<u>45</u>	<u>48</u>	<u>68</u>	<u>78</u>
	87	92	58	56	75	103
300 гПа	<u>59</u>	<u>59</u>	<u>63</u>	<u>63</u>	<u>72</u>	<u>75</u>
	95	86	65	69	121	100

1. Белов П.Н., Борисенков Е.П., Панин Б.Д. Численные методы прогноза погоды. Л.: Гидрометеиздат, 1989. 375 с.
2. Матвеев Л.Т. Курс общей метеорологии. Физика атмосферы. Л.: Гидрометеиздат, 1984. 751 с.
3. Гандин Л.С., Каган Р.Л. Статистические методы интерполяции метеорологических данных. Л.: Гидрометеиздат, 1976. 359 с.
4. Быков В.В., Курбаткин Г.П. // Изв. АН СССР. Сер. Геофиз. 1961. № 2. С. 307–318.
5. Руководство по прогнозированию метеорологических условий для авиации / Под ред. К.Г. Абрамович и А.А. Васильева. Л.: Гидрометеиздат, 1985. 301 с.
6. Справочник вертикальных корреляционных связей температуры, геопотенциала и плотности в свободной атмосфере над Северным полушарием. / Под ред. О.Б. Мерцалова. М.: НИИАК, 1966. 369 с.
7. Комаров В.С., Ременсон В.А. // Оптика атмосферы. 1988. Т. 1. № 7. С. 3–16.
8. Фаддеев Д.К., Фаддеева В.Н. Вычислительные методы линейной алгебры. М.: Физматгиз, 1963. 734 с.

Российский государственный гидрометеорологический университет,
Санкт-Петербург

Поступила в редакцию
28 апреля 1998 г.

V.I. Akselevich. Study of an Opportunity to Use the Autoregression Model for the Forecast of Temperature.

Results of forecasting the temperature at various levels and depending on various variants of the algorithm operation are considered. Comparison of results with inertial and climatic forecast is made. Optimum variants of the algorithm operation are found. The variants of initial data for the forecast and opportunities of obtaining an information about vertical profiles of temperature in various situations are discussed.