

А.С. МАЗМАНИШВИЛИ, д-р физ.-мат. наук, НТУ «ХПИ» (г. Харьков),
В.А. ПУЛЯЕВ, канд. техн. наук, Институт ионосферы (г. Харьков)

МЕТОД СТАТИСТИЧЕСКИХ ИСПЫТАНИЙ ЭФФЕКТИВНОСТИ АЛГОРИТМОВ ОЦЕНКИ ИОНОСФЕРНЫХ ПАРАМЕТРОВ В МЕТОДЕ НР

У статті приведено експериментування зі статистичною моделлю багатопараметричної системи з метою виявлення впливу методів обробки на точність і вірогідність оцінки іоносферних даних.

Numerical experimental results of with manyparametric statistical modelare given with the purpose of revealing influence of processing methods on accuracy and reliability of ratings ionospheric data.

1. Целостное представление о состоянии околоземной плазмы - ионосферы - может быть получено с помощью радарного метода некогерентного рассеяния (НР) [1]. Проблема состоит в том, что определение геофизических данных с его помощью исторически базируется на непараметрической статистике, когда физические эксперименты основываются на измерении сигнальных функций $f(x)$, образуемых из откликов ионосферы на зондирующие сигналы $x(t)$. Сигнальные функции тем или иным способом связаны с вектором ионосферных параметров $\vec{\theta}$, лишь последующее определение которых является целью проводимого опыта. Это означает, что в процессе наблюдений при ионосферных исследованиях вначале с использованием статистического усреднения производится оценивание некоторых функциональных характеристик ответного сигнала, например, высотно-временного распределения его автокорреляционных функций (АКФ). Последующий анализ подразумевает уже интерпретацию полученных АКФ для оценки по ним высотно-временного распределения тепловых флуктуаций электронной плотности и многих других физических параметров, определяющих ход наблюдаемых процессов в ионосфере.

2. Анализ последних публикаций показывает, что в настоящее время применяются специальные методы оценки корреляционных или спектральных данных, которые основаны на сравнении экспериментальных теоретических функций с полученными из модельных представлений об ионосфере. Обычно подобные методы реализуются на базе различных алгоритмов спуска [2]. Это довольно медленная и трудоемкая процедура вычислений, требующая дальнейшего усовершенствования и оптимизации вычислительного процесса, с помощью которого радар в автоматизированном режиме в реальном времени «on line» обеспечил бы функционирование наиболее быстрых и объективных алгоритмов определения геофизических параметров.

Известны решения данной проблемы, когда методология корреляционного анализа в методе НР использует элементы статистического подхода к вычислительным процедурам [3]. Но для создания системы автоматизированного анализа геофизических временных рядов требуется усовершенствование алгоритмов обработки и принятия решений с целью их оптимизации с точки зрения практически значимых критериев качества.

Суть параметрической идентификации в методе НР состоит в том, что исследуемый высотный диапазон условно разбивается на i интервалов с центрами на h_i . Внутри каждого интервала по отношению к вычисленной АКФ ставится в соответствие свой параметрический вектор $\bar{\theta}$, который в общем случае принимается равным [4]

$$\bar{\theta}(P') = \{N_e, T_i, T_e, f_m^+, \delta\}, \quad (1)$$

где i – число параметров в таком составе: N_e – электронная плотность плазмы; T_i и T_e – ионная и электронная температуры заряженных частиц; f_m^+ – относительные плотности для каждого из m вариантов ионного состава (комбинации кислорода, водорода, гелия, тяжелых ионов и др.); δ – доверительный интервал как функция статистического усреднения результатов.

На базе таких допущений с целью определения плазменных параметров при фиксации среднеквадратического отклонения $\chi^2(\bar{\theta})$ между экспериментальной и рядом модельных АКФ для параметрического вектора $\bar{\theta}$ отыскивается минимум вида

$$\chi^2(\bar{\theta}) = \frac{1}{k^2} \sum_{j=1}^k \left[\frac{r_{\text{изм}}(j\tau) - r_{\text{мод}}(j\tau, \bar{\theta})}{\sigma(j\tau)} \right]^2 \quad (2)$$

где индекс τ представляет задержку; $r_{\text{изм}}$ – измеренные реальные части АКФ для k вариантов задержек; $r_{\text{мод}}$ – модельная функция при тех же условиях; σ – расчетная неуверенность измерения, или вес ординат АКФ.

3. Так как вычислительные проблемы в значительной степени зависят от доступных ресурсов и развития техники, потребуем, чтобы качество данных было достаточным без затрат чрезвычайно больших усилий на получение значений соответствующих параметров с нереалистически высокой точностью. Для решения этой задачи для оценки эффективности функционирования алгоритмов в системе параметрической идентификации применим математическое моделирование процедуры идентификации, что позволит в дальнейшем решать задачи, связанные с построением оптимальных технических систем обработки.

4. Таким образом, целью настоящей работы является реализация на базе выражений (1) и (2) специальной статистической процедуры, с помощью которой возможно многократное повторение процесса идентификации параметров на каждом из высотных участков. Полученное в результате испыта-

ний множество случайных исходов решения задачи идентификации даст возможность, опираясь на методы математической статистики, определить искомые оценки для определения оптимального соотношения между точностью параметров вектора $\vec{\theta}$ и значениями их доверительных интервалов.

5. Предлагаемая математическая модель функционирования системы статистических испытаний изображена на рис. 1.

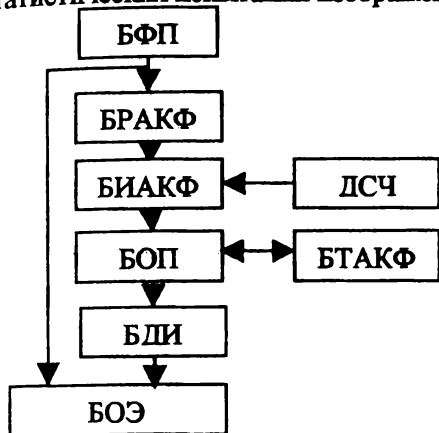


Рис. 1. Блок-схема системы статистических испытаний

Принцип действия модели следующий. Для проведения каждого испытания в блоке БОП по заданному параметрическому вектору $\vec{\theta}$, моделируется состояние плазмы заданием совокупности численных значений P'_i ее параметров, которые генерируются как некоторая реализация высотного изменения значений. Согласно этому параметрическому вектору в блоке БРАКФ вычисляются соответствующие формы экспериментальных АКФ при задаваемом количестве и местоположении k ординат $r_{\text{эксп}}$ АКФ.

С помощью датчика (в блоке ДСЧ) случайных чисел значения этих ординат в блоке БИАКФ подвергаются случайному искажению с помощью аддитивной помехи $n(t)$, что имитирует результат ее воздействия при использовании соответствующих технических средств с определенными точностными характеристиками. Под аддитивной помехой в нашей модели будем понимать суперпозицию случайной функции, не связанной с сигналом возбуждения. Это могут быть шумы аппаратуры или сигналы, порожденные внешними источниками, и, таким образом, введенный в модель параметрический шум следует рассматривать как учет неизбежного искажения формы АКФ, связанный со многими факторами. При зашумливании АКФ модель позволяет по линейному закону увеличивать интенсивность аддитивного шума по мере перебора ординат АКФ, и по экспоненциальному закону — дополнительно увеличивать его интенсивность по мере продвижения вдоль вы-

Модель состоит из следующих блоков: формирования параметров заданного вектора $\vec{\theta}$, ФП, расчета формы экспериментальной АКФ $r_{\text{эксп}}$ БРАКФ, искажения формы экспериментальной АКФ $r_{\text{эксп}}$ БИАКФ, оценки найденных параметров $\vec{\theta}$ по минимуму $\chi^2(\vec{\theta})$ БОП, вычисления доверительных интервалов $\Delta\theta$ БДИ, оценки показателя эффективности совпадения БОЭ, набора теоретических АКФ $r_{\text{теор}}$ БТАКФ, датчика случайных чисел ДСЧ.

соты.

Искаженные значения АКФ, представляющие наблюдаемую плазму в том виде, в каком ее воспринимает система, поступают на вход блока оценки БОП, в котором методом покоординатного спуска определяется принадлежность этой функции наиболее близкому по среднеквадратичному критерию модельному значению. Соответствующий этому значению определенный параметрический вектор $\bar{\theta}_o$ в блоке БДИ используется для вычисления доверительных интервалов по каждому параметру P_o^i согласно выражению

$$\Delta P_o^i = \bar{P}_o^i \pm \frac{1}{n^2} \sqrt{\sum_{k=1}^n (\bar{P}_o^i - P_{ok}^i)^2} \quad (3)$$

а в блоке БЭС – для сравнения результатов оценки $\bar{\theta}_o$ с "истинным" параметром $\bar{\theta}_*$, т.е. с тем, который задавался на этапе формирования АКФ.

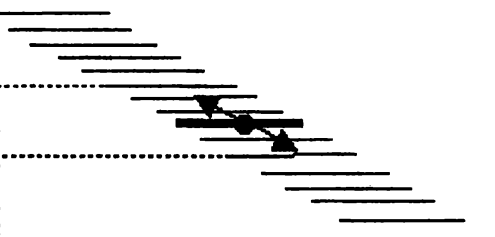
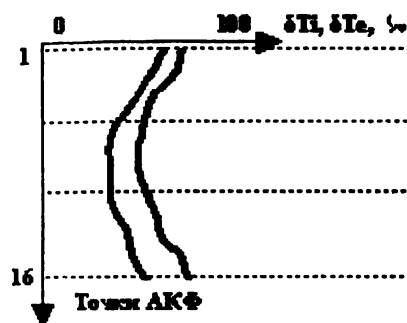
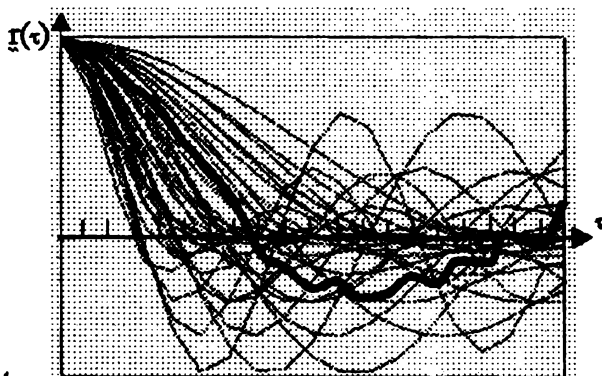
Число испытаний n определяется доверительной вероятностью, задаваемой при формулировке задачи исследования.

Рассмотренная статистическая модель позволяет найти зависимость показателя эффективности системы идентификации параметров в зависимости от вида и количества привлекаемых для распознавания признаков P_o^i . Результаты некоторого моделирования, в частности, при изменении T_i и T_e для ночной и дневной ионосферы представлены на рис.3. Приведен пример вычисления ~70-процентных доверительных интервалов для оценки кинетических температур в случае использования конечного ($k=20$) числа ординат АКФ. Ширина интервала δ лежит в пределах от единиц процентов (на высоте ~290 км) до десятков процентов (на высоте ~430км). Фрагменты слева дают представления о степени зашумленности экспериментальных АКФ при их моделировании для верхнего участка.

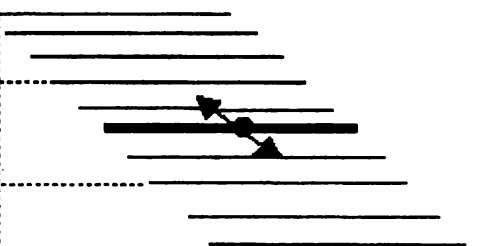
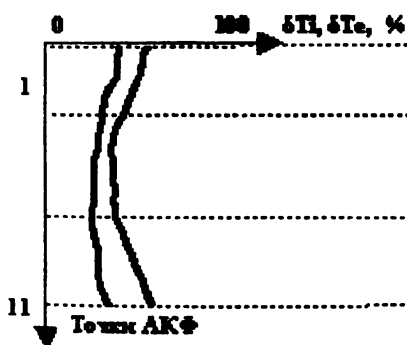
Приведем пример моделирования оценки точности совпадения исходных и найденных параметров и в случае варьирования местоположением и количеством k привлекаемых к идентификации ординат АКФ. Выбор их комбинаций имеет исключительно важное значение, поскольку связан с минимизацией затрат машинного времени, а также объемом записей в базе данных. Слишком малое k неизбежно влечет за собой потерю информации, снижает точность и достоверность статистических расчетов. Слишком большое k содержит определенную информационную избыточность. А поскольку все вычисления по статистическим алгоритмам имеют число шагов, прямо пропорциональное его величине, и повторяются в ПК до десятков тысяч раз, то имеется необходимость в поиске такого его значения, которое при достаточной точности расчетов было бы минимальным.

На рис. 2 приведены результаты поиска параметров по одной из среднезашумленной АКФ среди множества модельных в случае использования

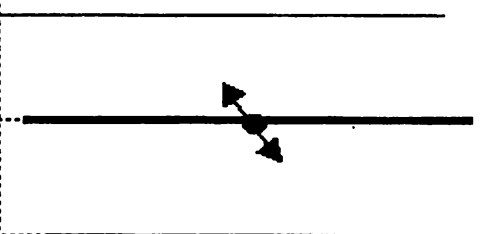
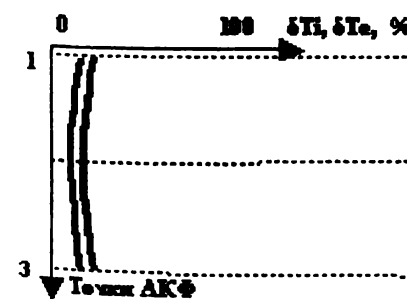
Рис.2. Влияние на точность
оценки параметров плазмы
количества и местоположения
используемых точек АКФ



а)



б)



в)

5 (а), 11 (б) и 18 (в) ординат и при их различных положениях. Оцениваемая ширина доверительного интервала в этом случае приняла значения от 10% до 100%, что может служить основой для выбора наиболее оптимального сочетания точности, времени вычислений и аппаратных ресурсов.

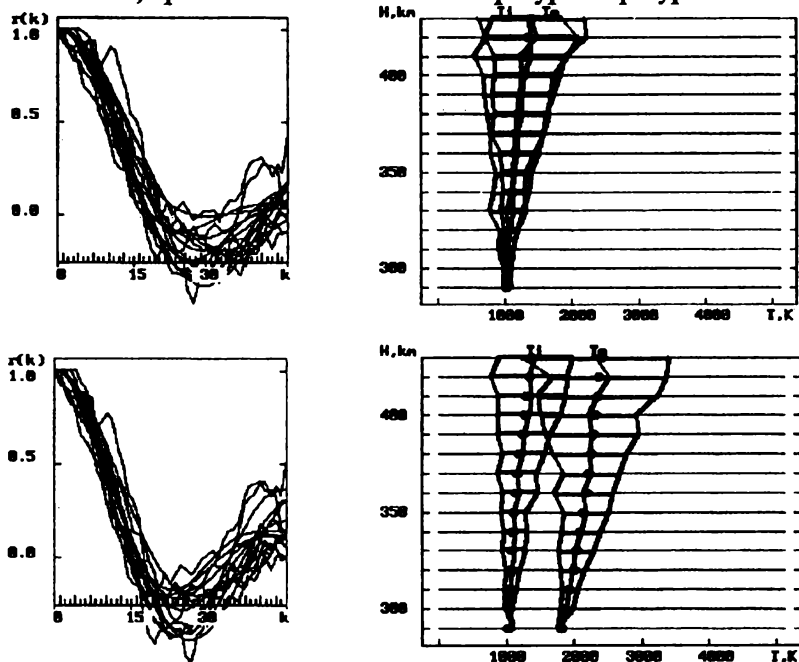


Рис.3. Доверительные интервалы при вычислении кинетических температур ионосферной плазмы для ночной и дневной ионосферы

6. Выводы. Экспериментирование с разработанной статистической моделью позволяет решать задачи построения оптимальных систем оценки ионосферных данных. Получаемые на базе модели сведения являются исходными при выборе сложности алгоритмов обработки, при определении состава технических средств наблюдения, при оптимальном распределении точностей по вычислительным средствам, которыми оснащается распознающая система.

Список литературы: 1. Таран В.И. Исследование ионосферы с помощью радаров некогерентного рассеяния в Харькове// Вестн. ХГПУ. Харьков, 1999. Вып.31. С.3-9. 2. Пуляев В.А. Статистическое оценивание параметров ионосферы в методе некогерентного рассеяния радиоволн. Радиотехніка: Воеукр. міжвід. наук.-техн. зб. 2002. Вип. 129. С.98-102. 3. Мазманишвили А.С., Рогожкін Е.В. Об эффекте смещения оценок параметров ионосферной плазмы при обработке данных некогерентного рассеяния. Радиотехніка: Воеукр. міжвід. наук.-техн. зб. 1999. Вип. 109. С.20-26. 4. Пуляев В.А. Вычислительные методы при обработке корреляционных функций сигнала некогерентного рассеяния// Вестн. ХГПУ: Сб. науч. тр. Харьков: ХГПУ, 2000. Вып.103. С.94-96.

Поступила в редакцию 08.04.03