КОНФЕРЕНЦИЯ D

ФИЗИКА ТРОПОСФЕРЫ

СЕМИПАРАМЕТРИЧЕСКИЕ ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ В ЗАДАЧАХ ОБРАБОТКИ РЕЗУЛЬТАТОВ МИНИ-СОДАРНЫХ ИЗМЕРЕНИЙ В АПС

В.А. Симахин,[1] Л.Г. Шаманаева[2,3]

[1]Курганский государственный университет

[2]Институт оптики атмосферы им. В.Е. Зуева СО РАН

[3]Национальный исследовательский Томский государственный университет

E-mail: sva_full@mail.ru, sima@iao.ru.

В докладе предложен метод нахождения семипараметрических робастных оценок максимального правдоподобия для семейства распределений Тьюки. Метод позволяет получать робастные оценки не только для удаленных, но и внутренних выбросов. Показано, что данные оценки сходятся к оценкам максимального правдоподобия в условиях полупараметрических задач, когда доля и распределение выбросов неизвестны. Полученные оценки использованы для обработки результатов мини-содарных измерений вертикальных профилей компонентов скорости ветра в АПС, в которых, присутствует значительное количество как удаленных, так и внутренних выбросов. На конкретных примерах показана высокая эффективность предложенных оценок.

Исследования результатов мини-содарных измерений скорости ветра в АПС показывает присутствие как удаленных, так и внутренних выбросов [1, 2]. Сложность проблемы связана с тем, что распределение и доля выбросов, как правило, неизвестны. Задачи такого типа относятся к классам семипараметрических или семинепараметрических задач [2, 3]. Стандартные методы обработки данных измерений в этом случае приводят к низкой эффективности и значительным смещениям оценок различных характеристик компонентов скорости ветра, и требуется применение робастных и непараметрических методов статистики [2–5]. Типовые методы и процедуры робастной статистики, позволяющие снизить влияние выбросов, связаны с жестким или мягким усечением выделяющихся наблюдений [4, 5]. При таком подходе возникают проблемы при определении адаптивных робастных процедур построения оценок, существенно снижающих вес и влияние как удаленных, так и внутренних «не выделяющихся» выбросов. Отметим, что внутренние выбросы на уровне $(1–2)\sigma$ могут существенно изменять как форму распределения, так и вносить неконтролируемые смещения в оценки параметров. Рассмотрим построение адаптивных робастных оценок на семипараметрических классах распределений Тьюки. Пусть $\vec{X}_N = (x_1,...,x_N)$ – выборка н.о.р. с функцией распределения (ф.р.) $F(\vec{t},\theta) \in \mathrm{P}_\theta$ – супермодель Тьюки

$$\mathrm{P}_{\theta}=\{F(x,\theta)=(1-\varepsilon)G(x,\theta)+\varepsilon\cdot H(x,\theta)\}, \tag{1}$$

где $G(\vec{t},\theta)\in \mathrm{P}_{0\theta}$ – априорная модель ф.р., $H(x)$ – распределение выбросов, $\varepsilon\geq 0$ – доля выбросов. Требуется по выборке $\vec{X}_N$ оценить неизвестный параметр θ.

Супермодель Тьюки используется в качестве удобной модели реальных распределений $F(\vec{t},\theta)$, которые можно считать приближенно совпадающими с априорным распределением $G(\vec{t},\theta)$ [5]. Обозначим через $f(x), g(x), h(x)$ плотности распределений F, G, H, соответственно и $F_N(x)$ – эмпирическую функцию распределения (э.ф.р.). Для модели Тьюки (1) оценка максимального правдоподобия (ОМП) θ_N параметра θ определяется из оценочного уравнения

$$\int\varphi(x,\theta_N)dF_N(x)=\frac{1}{N}\sum_{i=1}^{N}\varphi(x_i,\theta_N)=0\,, \tag{2}$$

где оценочная функция $\varphi(x,\theta)$ представима в виде

$$\varphi(x,\theta)=\frac{\partial}{\partial\theta}\ln g(x,\theta)\cdot\frac{g(x,\theta)}{f(x,\theta)}=U(x,\theta)\cdot W(x,\theta), \tag{3}$$

$U(x,\theta)$ – функция вклада ОМП для $G(\vec{t},\theta)$, $W(x,\theta)$ – весовая функция

$$W(x,\theta)=\frac{(1-\varepsilon)g(x,\theta)}{f(x,\theta)}=\left[1+\frac{\varepsilon\cdot h(x)}{(1-\varepsilon)\cdot g(x,\theta)}\right]^{-1}. \tag{4}$$

Анализ выражения (4) показывает, что в нашем случае эффективная оценка представляет взвешенную ОМП априорного распределения $G(\vec{t},\theta)$ с весами $W(x,\theta)$ (4), которые и определяют расхождение между $G(\vec{t},\theta)$ и $\{H(x),\varepsilon\}$. В параметрическом случае, когда вид $G(\vec{t},\theta)$ и $\{H(x),\varepsilon\}$ известен, в соответствии с (2)–(4) можно синтезировать эффективную взвешенную ОМП – параметрическая задача. В задачах классической робастной статистики информация о выбросах $\{H(x),\varepsilon\}$ неизвестна – семипараметрическая задача [3, 4, 5]. Именно семипараметрические задачи и доставляют основные проблемы при синтезе робастных оценок [4, 5]. В дальнейшем предполагается, что информация о $\{H(x),\varepsilon\}$ неизвестна – семипараметрическая задача [3, 4]. Так как информация о $\{H(x),\varepsilon\}$ неизвестна, то неизвестна и плотность распределения $f(x,\theta)$. Для определения оценки ОМП определим оценочное уравнение (2) в виде

$$\int\varphi_N(x,\theta_N)dF_N(x)=\frac{1}{N}\sum_{i=1}^{N}\varphi_N(x_i,\theta_N)=0\,, \tag{5}$$

где оценочная функция $\varphi_N(x,\theta)$ определена в виде

$$\varphi_N(x,\theta)=\frac{\partial}{\partial\theta}\ln g(x,\theta)\cdot\frac{g(x,\theta)}{f_N(x,\theta)}=U(x,\theta)\cdot W_N(x,\theta)\,, \tag{6}$$

$U(x,\theta)$ – функция вклада ОМП для $G(\check{t},\theta)$, $W_N(x,\theta)$ – весовая функция

$$W_N(x,\theta)=\frac{(1-\varepsilon)g(x,\theta)}{f_N(x,\theta)},$$

$$f_N(x)=\frac{1}{N\sqrt{2\pi}\cdot h_N}\sum_{i=1}^{N}\exp\left(-\frac{1}{2}\left(\frac{x-x_i}{h_N}\right)^2\right) \tag{7}$$

– непараметрическая оценка плотности Розенблата-Парзена. Например, для нормального распределения оценочное уравнение (5) имеет следующий вид

$$\sum_{i=1}^{n}(x_i-\theta)\frac{\frac{1}{\sigma\sqrt{2\pi}}\exp(-\frac{(x_i-\theta)^2}{2\sigma^2})}{f_N(x_i)}=0\,. \tag{8}$$

Полученные оценки (5)–(8) использовались для обработки данных допплеровского мини-содара AV4000. Рабочая частота содара 4900 Гц, длительность импульса излучения 60 мс, период повторения импульсов 4 с. Излучение последовательно посылалось в трех направлениях – вертикально вверх и под углами 14° к вертикали в двух взаимно ортогональных плоскостях. Анализировались данные измерений трех компонентов скорости ветра в 43 высотных стробах вертикальной протяженностью 5 м в диапазоне высот 5–200м. Обрабатывались серии из N = 150 профилей, что обеспечивало усреднение за 10-минутный период измерения.

На рисунке 1 приведены классические оценки (пунктирные кривые) и полупараметрические оценки (сплошные кривые) высотного профиля x-компоненты скорости ветра и его дисперсии, построенные по результатам мини-содарных измерений в утренние часы с 08:00 до 08:10 местного времени, которые подтверждают эффективность предложенного алгоритма. Видно, что предложенный алгоритм обеспечивает существенно меньшие значения дисперсии оценки на больших высотах. Предложенный метод был использован для анализа суточных вариаций средних значений и дисперсий высотных профилей x, y, z компонентов скорости ветра в АПС по результатам мини-содарных измерений в утренние (с 08:00 до 08:10 местного времени), дневные (с 14:00 до 14:10), вечерние (с 19:00 до 19:10), и ночные часы (с 00:00 до 00:10) . На рисунке 2 представлены семипараметрические оценки суточных вариаций вертикальных профилей x-компонента скорости ветра и его дисперсии.

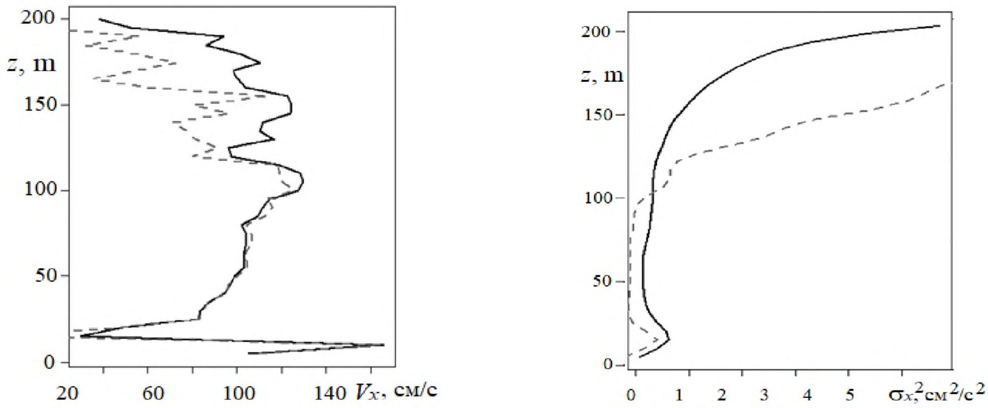

Рисунок 1 – Высотные профили *x*-компонента скорости ветра и его дисперсии: пунктирные кривые – классические оценки; сплошные кривые – полупараметрические оценки (4)–(7)

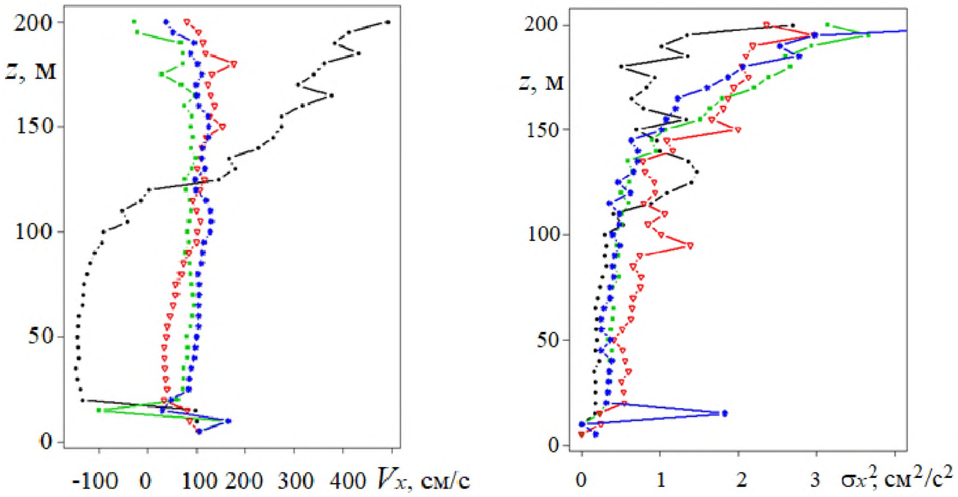

Рисунок 2 — Полупараметрические оценки суточных вариаций вертикальных профилей *x*-компонента скорости ветра и его дисперсий в утренние (синие кривые), дневные (зеленые кривые), вечерние (красные кривые), и ночные часы (черные кривые)

Результаты обработки мини-содарных измерений в АПС показывают, что стандартные методы обработки приводят к значительным смещениям и низкой эффективности оценок по сравнению с семипараметрическими оценками максимального правдоподобия. Работа выполнена в рамках государственного задания II.12.1.2. ИОА СО РАН (рег. № проекта АААА-А17-117021310152-4).

1. *Федоров В.А.* Измерение содаром "Волна-3" параметров радиальных компонент вектора скорости ветра // Оптика атмосферы и океана. 2003. Т. 16. № 02. С. 151-155.

2. *Симахин В.А., Черепанов О.С., Шаманаева Л.Г.* Пространственно-временная динамика скорости ветра по результатам мини-содарных измерений // Известия вузов. Физика. 2015. Т. 58. № 12. С. 176–181.

3. *Симахин В А.* Адаптивные оценки. Курган: КГУ, 2019. 240 с.

4. *Шуленин В.П.* Робастные методы математической статистики. – Томск: Изд-во НТЛ, 2016. 260 с.

5. *Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В.* Робастность в статистике. – М.: Мир, 1989. 512 с.