

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(СПбГУ)
Кафедра физики атмосферы
Направление «Прикладная математика и физика»



ИССЛЕДОВАНИЕ ВОДОЗАПАСА ОБЛАКОВ МКВ МЕТОДОМ В РАЙОНЕ САНКТ-ПЕТЕРБУРГА

Магистерская диссертация студента:
Бирюкова Егора Юрьевича

Научный руководитель:
д. ф.-м. н., в.н.с. **Косцов В.С.**

Рецензент:
д. ф.-м. н., с.н.с. **Заболотских Е.В.**

Санкт-Петербург
2017

Оглавление

1. Введение	3
2. Методы измерения водозапаса облаков	4
3. Примеры МКВ радиометрических измерений водозапаса	6
4. Постановка задачи	10
4.1. Описание основных проблем и формулировка задачи	10
4.2. Описание МКВ радиометра RPG НАТPRO	15
4.3. Описание процедуры нахождения коэффициентов МЛР и вычисления значений водозапаса облаков	16
4.4. Статистическая модель атмосферы и множественное моделирование измерений	19
5. Исследование погрешностей определения водозапаса алгоритмом МЛР на основе модельных расчетов	21
6. Применение множественной линейной регрессии к данным, полученным МКВ радиометром RPG НАТPRO	29
6.1. Исследование результатов, полученных МЛР в периоды с большой вариацией систематической составляющей погрешности определения водозапаса облаков	29
6.2. Сопоставление результатов МЛР с результатами, полученными другими алгоритмами, для облачности с большим и малым водозапасом	33
7. Основные результаты и выводы	39
Список литературы	41

1. Введение

Облака являются важным составляющим элементом формирования погоды и климата планеты. Облака оказывают влияние на такие процессы, как: круговорот воды в природе, тепловой баланс системы Земля – атмосфера, самоочищение атмосферы, прохождение электромагнитных волн в атмосфере и др. В последние десятилетия сведения о характеристиках облаков стали остро востребованы в областях прогнозирования погоды, изучения и моделирования климата, активных воздействий на атмосферные процессы, космических исследований, а также авиационной и радиолокационной техники, спутниковой метеорологии и т.д.[1]

Одной из основных характеристик облака является его водозапас. Согласно справочнику «Облака и облачная атмосфера» 1989 года [1], водозапас – масса сконденсированной воды в столбе облачного воздуха единичного сечения. Измерение водозапаса облаков является одним из важнейших звеньев в изучении строения, эволюции и физических свойств естественных облаков. Для получения водозапаса облаков используются различные активные и пассивные дистанционные методы измерений. Например, МКВ – радиометрический метод. Данный метод основан на регистрации собственного радиотеплового излучения и является одним из немногих эффективных методов определения водозапаса облаков с поверхности Земли, ввиду возможности осуществления длительных, автоматических измерений данным методом в практически любых погодных условиях с высоким временным разрешением. МКВ метод широко используется для наземных измерений, а также для спутниковых измерений над океаническими поверхностями.

Данная диссертационная работа состоит из двух частей:

- Обзор методов для измерения водозапаса облаков и соответствующей аппаратуры (МКВ радиометров);

- Реализация метода множественной линейной регрессии для обработки данных микроволнового радиометра RPG НАТPRO, функционирующего в СПбГУ.

2. Методы измерения водозапаса облаков

Все способы измерения атмосферных параметров имеют общую классификацию. Прежде всего, все методы подразделяются на:

- 1) Контактные методы измерений
- 2) Дистанционные методы измерений

Контактные методы измерений подразумевают измерения атмосферного параметра непосредственно в определенном объеме или площади атмосферы. Существуют приборы для измерения облачных параметров (в частности влажности и размера капель) контактным методом измерений, например прибор СИВ А-27, устанавливаемый на борту самолета[2]. Подобное оборудование обладает хорошей точностью (погрешность для СИВ А-27 составляет 10-30%) и большим количеством измерений, но очень малой областью измерения (1-2 измерения в облаке), что делает невозможным измерение интегральной влажности или водозапаса, а также временных и пространственных распределений данных величин. К тому же контакт прибора с измеряемой средой может повлиять на сами измерения (например, самолет, пролетающий через облако, может нарушить его структуру). Еще один недостаток подобных методов – их дороговизна (необходима аренда самолета). Поэтому контактные методы измерений не подходят для получения информации о состоянии атмосферы в региональном или глобальном масштабах.

Недостатки контактных методов измерений призваны решить дистанционные методы. Дистанционные методы измерений в общем случае основаны на регистрации различных полей – гравитационных, электрических, магнитных, электромагнитных, акустических. Измерения происходят на расстоянии от исследуемого объема или площади

поверхности, причем это расстояние может быть значительным (например, для спутниковых измерений). Дистанционные методы измерений имеют свою классификацию [3]. В первую очередь они делятся на активные и пассивные. Пассивные методы используют измерения характеристик естественных полей излучения и делятся на:

- 1) Методы собственного излучения атмосферы
- 2) Методы рассеянного излучения
- 3) Методы ослабления и поглощения (прозрачности) излучения
- 4) Методы рефракции
- 5) Методы отраженного излучения

Активными методами зондирования, использующими искусственные источники электромагнитного излучения, называются:

- Лазерное зондирование
- Радиолокационное зондирование
- метод рефракции
- метод ослабления и поглощения

Также, дистанционные методы делятся по расположению приемника на:

- Наземные методы
- Спутниковые методы
- Измерения с наземного, морского и воздушного транспорта

Из вышперечисленных методов измерений атмосферных параметров следует выделить несколько методов, применимых к измерению водозапаса облаков. Из активных методов – радиолокационное и лидарно-радиометрическое зондирование. Физические принципы данных методов достаточно просты: в направлении измеряемого объекта посылаются импульсы электромагнитного излучения определенных частот (радиоволновой диапазон для радиолокации и от видимого до далекого ИК излучения для лидарного зондирования), после чего приемник системы регистрирует отраженные (рассеянные назад) от объекта сигналы. В

частности, лидарно-радиометрический метод зондирования описан в диссертации Захаровой П.В.: «Лидарно-радиометрические методы контроля водности и водозапаса жидкокапельных облаков» [2], в которой можно найти:

- Методику численных измерений водозапаса
- Используемую аппаратуру и ее калибровку
- Экспериментальные данные лидарно-радиометрического зондирования в Москве за июль-август 1993-1994 года (зависимости водозапаса от излучательной способности, водности и т.д.)

Из пассивных дистанционных методов необходимо выделить МКВ-радиометрический метод измерений, основанный на регистрации собственного радиотеплового излучения атмосферы. Определение водозапаса атмосферы методом МКВ радиометрии основано на их связи с характеристиками радиотеплового излучения атмосферы: радиояркостной температурой и оптической толщой. Эти связи следуют из уравнения радиотеплового излучения атмосферы:

$$I_{\nu} = \frac{2k\nu^2}{c^2} \int_0^{\infty} T(l)\gamma_{\nu}(l) \times \exp\left(-\int_0^l \gamma_{\nu}(l')dl'\right)dl + I_{\Phi} \exp\left(-\int_0^{\infty} \gamma_{\nu}(l)dl\right),$$

где I_{ν} – интенсивность излучения; γ_{ν} – коэффициент поглощения; ν – частота; dl – элемент пути; k – постоянная Больцмана; $T(l)$ – профиль термодинамической температуры атмосферы; I_{Φ} – интенсивность реликтового излучения космоса.

3. Примеры МКВ радиометрических измерений водозапаса

В качестве примера наземных МКВ радиометрических измерений приведем результаты исследования водозапаса облаков в г. Долгопрудный, Московская область радиометрическим комплексом «Микрорадком» [4]. В этой работе, в частности, показаны таблицы изменчивости водозапаса для разных периодов времени в 2012-2013 году и при разной температуре для

высококучевых и высокослоистых облаков (Табл. 1), а также при наличии одной лишь дымки (Табл. 2).

Табл.1. Водозапас для высококучевых и высокослоистых облаков

Водозапас для высококучевых и высокослоистых облаков				
Дата	$T, ^\circ\text{C}$	Время	Тип облачности	Водозапас, г/см ²
05.12.2012	-4,1	17:19	Высококучевые просвечивающие	0,0002–0,02
12.12	-9,7	22:24	Высококучевые просвечивающие	0,0022–0,02
22.12	-14,1	13:16	Высококучевые просвечивающие	0,01–0,04
23.12	-20,1	04:07	Высококучевые просвечивающие	0,02–0,03
22.01.2013	-14	7, 13, 22:24	Высококучевые просвечивающие	0,03–0,05
23.01	-11	00:04	Высококучевые просвечивающие	0,001–0,016
15.02	-2,6	7, 19	Высококучевые просвечивающие	0,01–0,05
30.03	1,6	8:10	Высококучевые просвечивающие	0,01–0,02
30.03	1,6	17:19	Высокослоистые непросвечивающие	0,12–0,13
04.04	5,1	10:13	Высококучевые просвечивающие	0,01–0,02
04.04	5,1	22:24	Высококучевые просвечивающие	0,01–0,05
23.04	11,3	11:13	Высокослоистые непросвечивающие	0,01–0,03
12.05	26,9	22:24	Высококучевые просвечивающие	0,00–0,11
17.07	20,8	04:10	Высококучевые просвечивающие	0,00–0,011
13.10	12,2	00:10	Высококучевые просвечивающие	0,01–0,02
13.10	12,2	22:24	Высококучевые непросвечивающие	0,01–0,04
23.10	1,8	08:24	Высококучевые просвечивающие	0,03–0,06
29.10	14	13:16	Высококучевые просвечивающие	0,03–0,05
30.10	11,1	03:04	Высококучевые просвечивающие	0,02–0,052
18.11	2,6	00:04	Высококучевые просвечивающие	0,4–0,6

Табл.2. Водозапас при отсутствии облаков и наличии дымки

Водозапас при отсутствии облаков и наличии дымки			
Дата	T, °C	Время	Водозапас, г/см ²
20.12.2012	-17,5	00:07	0,0016–0,0047
22.12	-14,1	00:04	0,005–0,007
23.12	-20,1	19:24	00–0,015
15.02.2013	-2,6	22:24	0,01–0,017
16.02	-6,3	00:07	0–0,01
08.04	8,1	04:10	0,0004–0,011
13.04	8,8	04:07	00–0,007
25.06	24,4	00:07	00–0,001
29.06	20,3	00:10	0,02–0,03
13.07	24,1	00:04	0,03–0,05
10.08	28,3	00:01	0,0009–0,01
10.08	29,2	19:24	0,02–0,04
17.08	21	00:10	0,02–0,03
27.08	15,2	00:04	00–0,02
19.09	13,1	00:04	0,0009–0,001
27.09	7,6	08:10	0,002–0,006
23.10	1,8	00:07	0,002–0,012
09.11	7,5	07:13	0,004–0,026

Также, измерения водозапаса проводились в п. Воейково, Ленинградской области МКВ радиометрами при частотах зондирования 9.3ГГц и 36.5ГГц. В приведенных результатах анализа измерений водозапаса [5] говорится: в холодный период в 95% времени водозапас облаков в зените не превышал 0.28 кг/м², а значения водозапаса облаков, превышающие 0.52 кг/м² регистрировались лишь в 1% случаев, в теплый период года лишь в 5% случаев водозапас облаков превышал 0.7 кг/м²..

Экспериментальные исследования атмосферы с помощью наземной микроволновой аппаратуры начались на физическом факультете СПбГУ в 2007 году. В рамках экспериментов осуществлялись дистанционные измерения концентрации озона в стратосфере и нижней мезосфере с помощью радиометра, разработанного Институтом прикладной физики РАН (ИПФ РАН, Нижний Новгород) и работающего в области линии поглощения озона 110 ГГц [6]. С 2012 года был введен в эксплуатацию микроволновый радиометр RPG-НАТPRO производства немецкой компании Radiometer Physics GmbH [7].

Помимо вышеописанных, измерения проводятся также с помощью МКВ радиометров: MICCY (MICrowave radiometer for Cloud CartographY) [8],

Radiometrics MP-3000A [9], ASMUWARA (All-Sky MUlti Wavelength RAdiometer) [10], TEMPERA (TEMPERature RAdiometer) [11].

Для примера спутниковых МКВ радиометрических измерений водозапаса облаков можно привести данные с ресурса Climate Data Guide [12], которые демонстрируют распределение общего водозапаса и его погрешности за июль 1998 по океанам (Рис.1), а также распределение водозапаса на долготе 186Е и широте 5,5N (Рис.2). По рис. 1 видно, что основные значения водозапаса 50 – 100 г/м², и его погрешность 8 – 16 г/м². Рис. 2 выделяет из предыдущего рисунка изменчивость водозапаса по широте и долготе.

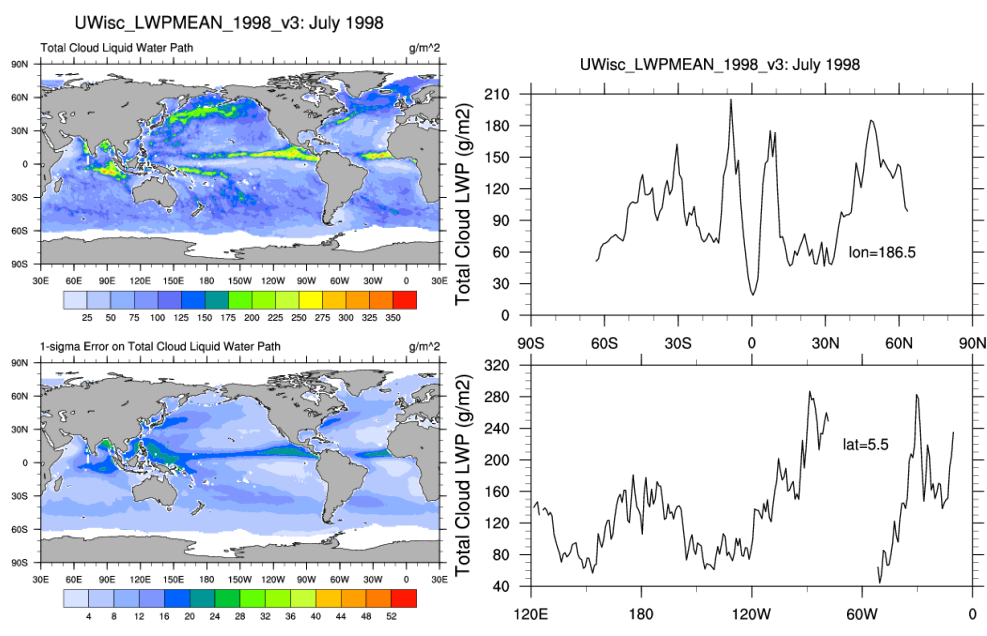


Рис 1. Водозапас и его рассчитанная погрешность (июль 1998 г.)

Рис 2. Водозапас для долготы 186Е и 120-150Е, а также широты 5,5N

Говоря о спутниковых МКВ измерениях, необходимо упомянуть интернет ресурс Remote Sensing Systems [13], где, помимо прочих измерений, представлены данные измерений водозапаса облаков над океанической поверхностью с помощью спутниковых МКВ радиометров: SSM/I [14], WindSat [15], AMSRE и AMSR2 [16], TMI [17]. Пример расчета водозапаса облаков с помощью данных AMSR2 можно увидеть на Рис. 3:

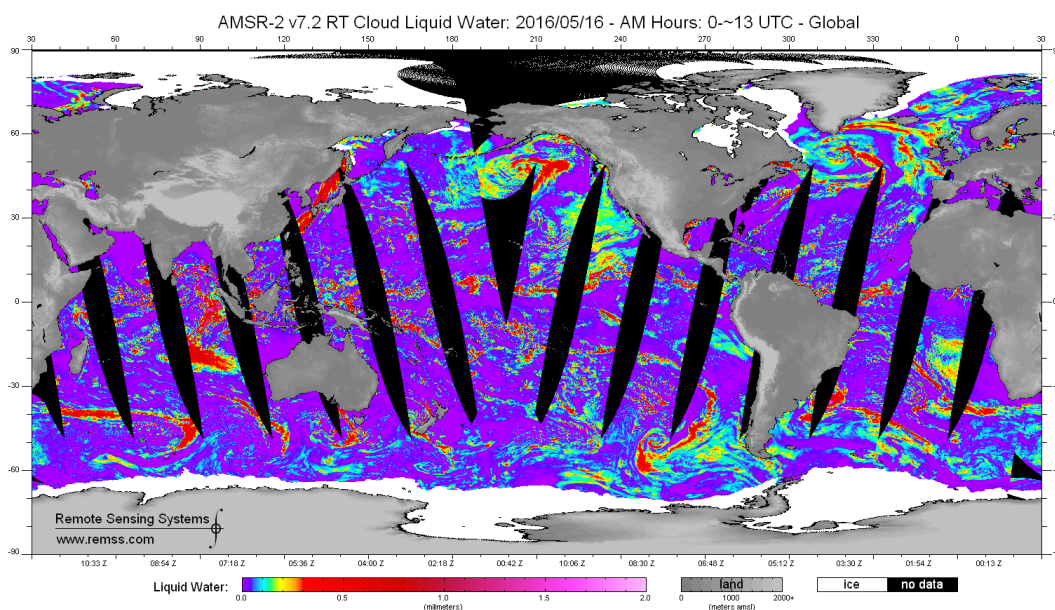


Рис. 3. Спутниковые измерения водозапаса облаков прибором AMSR2 для 16 Мая 2016 года.

4. Постановка задачи

4.1. Описание основных проблем и формулировка задачи

МКВ радиометр RPG НАТПРО, функционирующий на физическом факультете Санкт-Петербургского государственного университета, предназначен для решения следующих задач в автоматическом непрерывном режиме: определения профилей температуры и влажности в тропосфере, интегрального содержания водяного пара, водозапаса облаков. Определение данных величин осуществляется по измерениям радиояростной температуры МКВ излучения в 7 каналах в области полосы поглощения кислорода 0.5 см и 7 каналах в области линии поглощения водяного пара 1.35 см [7]. Измерения проводятся в постоянном режиме с временным шагом 1-2 секунды. Для получения данных по всем 4 задачам используются два отдельных, независимых алгоритма: множественная квадратическая регрессия, предоставленная разработчиком прибора и интегрированная в компьютерную программу управления, и, так называемый, физический алгоритм, основанный на обращении уравнения переноса излучения. Физический алгоритм был разработан в СПбГУ специально для МКВ

радиометра RPG HATPRO. В настоящей диссертационной работе рассматривается только задача определения водозапаса облаков. При определении значений водозапаса облаков прибором RPG HATPRO существует несколько проблем:

- проблема валидации полученных значений водозапаса облаков;
- проблема оценки систематической составляющей погрешности определения водозапаса облаков;
- Проблема определения качества полученных данных и исключения заведомо ложных результатов.

Способов валидации и оценки точности определения водозапаса облаков, полученных МКВ измерениями, немного. Одним из таких методов является анализ безоблачных периодов измерений, когда значения водозапаса облаков должны быть равны нулю. Данный метод был использован в работе [18], где проанализирована точность определения значений водозапаса облаков на данных, полученных двухканальным МКВ радиометром (23,8 и 31,4 ГГц). Также, в этой работе предложена комбинация статистических и физических алгоритмов, направленная на улучшение качества данных по водозапасу облаков, получаемых при измерениях двухканальным МКВ радиометром. Проблемы оценки систематической составляющей погрешности определения водозапаса облаков и исключения заведомо ложных результатов во время дождя и после дождя были рассмотрены в статье [19]: в данной статье было подчеркнuto, что проблема учета дождя в измерениях критична для анализа данных, потому что во время дождя на приборе скапливается вода и, в таком случае, измерения прибора перестают иметь физическую ценность до момента полного испарения воды на приборе. Поэтому, в данной статье было указано, что в общем случае из полученных данных необходимо исключить значения водозапаса облаков, полученные в течение 1 часа после дождя. Для того, чтобы получить оценку систематической составляющей погрешности определения водозапаса облаков, происходящей из-за сбоя прибора и из-за изменения его

характеристик со временем, авторы работы [19] выделяли безоблачные периоды измерений и использовали специальную программу, чтобы оценить, как повлияло исключение систематической составляющей погрешности определения водозапаса облаков из полученных данных. В работе [20] также описана проблема наличия шума в МКВ измерениях, вызванного осаждением воды как из-за увлажнения, так и из-за наличия осадков. Было подчеркнуто, что основной необходимостью является точное определение моментов времени начала осаждения и полного испарения влаги с поверхности обтекателя. Авторы работы [20] заметили почти полную невозможность определения периода времени полного испарения влаги для определенного радиометра без использования одновременных измерений с других приборов. В работе [21] указано, что оценку точности определения водозапаса облаков сложно осуществить из-за недостатка достоверных и независимых измерений, и, таким образом, полная оценка точности определения водозапаса должна осуществляться теоретически при учете всех возможных источников ошибок. Авторы работы [21] представили подробный анализ точности определения водозапаса облаков и предложили несколько способов ее улучшения. Подробный анализ природы погрешностей определения водозапаса облаков для двухканального радиометра TROWARA, был проведен в статье [22], и было выведено, что систематическая составляющая погрешности определения водозапаса облаков связана с изменчивым влиянием водяного пара в линии поглощения 31 ГГц. Также, в работе [22] большое внимание уделялось проблеме определения и анализа безоблачных периодов измерений.

Исходя из всего вышеперечисленного и сравнения результатов определения водозапаса регрессионным и физическим методом [23], можно сказать, что при применении регрессионного метода, созданного разработчиками аппаратуры и использующего квадратичную регрессию, возникают три более конкретные существенные проблемы:

- 1) Невозможность идентификации безоблачных ситуаций по критерию минимума вариаций водозапаса;
- 2) Значительные вариации систематической составляющей погрешности определения водозапаса.
- 3) Существенный (по сравнению с физическим методом) отклик на сбои во входных данных.

Особо остро, в частности, стоит проблема значительной вариации систематической составляющей погрешности определения водозапаса облаков. В ходе определения среднечасовых значений водозапаса облаков, полученных алгоритмом множественной квадратической регрессии, выяснилось, что систематическая составляющая погрешности определения водозапаса облаков для данного метода изменяется случайным образом. Более того, в некоторые периоды измерений величина вариации систематической составляющей достигает больших значений – более $0,1 \text{ кг/м}^2$. Один из таких периодов представлен на Рис.4 и Рис.5.

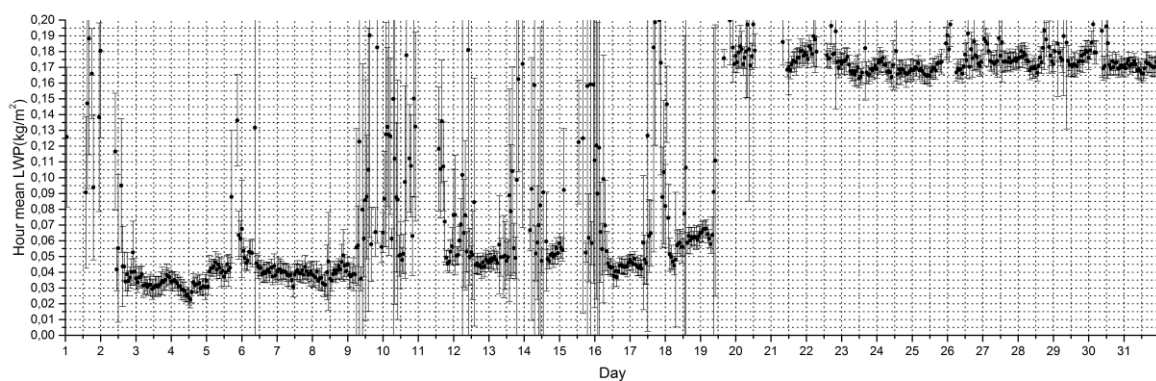


Рис.4. Среднечасовой ход водозапаса облаков в Августе 2013 года.

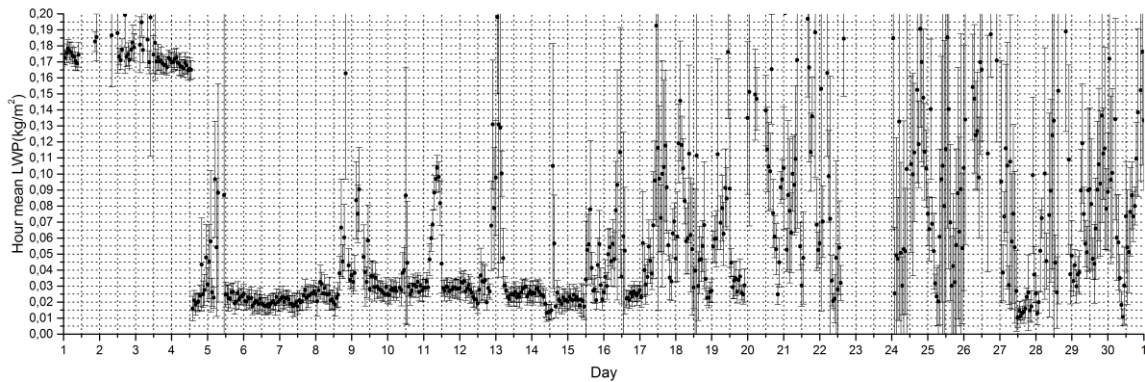


Рис. 5 Среднечасовой ход водозапаса облаков в Сентябре 2013 года.

Данное обстоятельство делает затруднительным дальнейший анализ полученных данных.

Настоящая диссертационная работа является продолжением исследований, описанных в статье [23] и ставит перед собой следующие цели:

- реализовать регрессионный метод определения значений водозапаса облаков, использующий линейное регрессионное соотношение и адаптировать его к данным измерений радиометра RPG-НАТPRO;
- исследовать погрешности линейного регрессионного метода на основе проведения численных экспериментов по определению водозапаса облаков;
- исследовать отклик линейного регрессионного метода на сбои во входных данных и оценить систематическую составляющую погрешности определения водозапаса облаков;
- провести сравнение результатов определения водозапаса облаков физическим методом с результатами, полученными регрессионными методами, использующими квадратическое и линейное регрессионное соотношение.

4.2. Описание МКВ радиометра RPG НАТPRO

Технические характеристики микроволнового прибора RPG-НАТPRO самым подробным образом представлены на сайте разработчика, а также доступны для загрузки руководства по установке, эксплуатации и проведению измерений [7]. В Табл. 3 представлены центральные частоты и значения полуширины так называемых «температурных» и «влажностных» каналов радиометра RPG-НАТPRO.

Табл. 3. Центральные частоты f и полуширина Δf «температурных» и «влажностных» каналов аппаратуры RPG-НАТPRO.

«влажностные»	1	2	3	4	5	6	7
f [GHz]	22.24	23.04	23.84	25.44	26.24	27.84	31.4
Δf [MHz]	230	230	230	230	230	23	230
«температурные»	8	9	10	11	12	13	14
f [GHz]	51.26	52.28	53.86	54.94	56.66	57.30	58.00
Δf [MHz]	230	230	230	230	600	1000	2000

Прибор установлен на вышке на крыше здания НИИ физики СПбГУ на базе научного парка СПбГУ в ресурсном центре «Геомодель», см. Рис.6:



Рис. 6. Внешний вид RPG-НАТPRO

Географические координаты прибора - 59.88107°N, 29.82597°E. Прибор установлен на 56 метров над уровнем моря. Угловое вертикальное

сканирование осуществляется в плоскости север-юг в направлении на север. Прибор работает непрерывно, измерения осуществляются в двух режимах. В основном режиме измеряется интенсивность нисходящего МКВ излучения в зенитном направлении – выдача данных 1 раз в 2 секунды. Дополнительный режим включается периодически (1 раз в 20 минут), в этом режиме производится угловое сканирование. Среднеквадратическая погрешность прибора при определении радиояркостной температуры – 0,2 К.

4.3. Описание процедуры нахождения коэффициентов МЛР и вычисления значений водозапаса облаков

В данном параграфе описана процедура нахождения коэффициентов множественной линейной регрессии для вычисления водозапаса облаков по радиояркостным температурам, полученным в разных частотных каналах МКВ радиометра RPG НАТPRO.

Пусть имеется некоторый обучающий массив, который мы будем называть «обучающим», с известными значениями водозапаса облаков и соответствующими им наборами радиояркостных температур.

Множественная линейная регрессия:

$$W(i) = \left(\sum_{j=1}^k x(j) * T_{\text{р}}(i, j) \right) + b + \varepsilon(i), \quad (1)$$

где $W(i)$ и $T_{\text{р}}(i, j)$ – значения водозапаса облаков (искомые значения) и радиояркостной температуры канала j (предикторы) для измерения i из выше обозначенного массива данных; $x(j)$ – коэффициенты МЛР; b – константа; ε - случайная ошибка; k – количество частотных каналов для измерения радиояркостной температуры. Поскольку при решении обратных задач дистанционного зондирования в большом числе случаев используют линейные приближения и работают не с самими величинами, а с их отклонениями от средних значений, в настоящей работе для сохранения

общности с другими методами мы также вместо $W(i)$ и $T_{\text{я}}(i, j)$ используем $\Delta W(i)$ и $\Delta T_{\text{я}}(i, j)$ – отклонения данных величин от их средних значений $W_{\text{ср}}$ и $T_{\text{яср}}(j)$, вычисленных по всем рассматриваемым измерениям:

$$\Delta W(i) = W(i) - W_{\text{ср}}, \quad (2)$$

$$\Delta T_{\text{я}}(i, j) = T_{\text{я}}(i, j) - T_{\text{яср}}(j), \quad (3)$$

$$\Delta W(i) = \left(\sum_{j=1}^k x(j) * \Delta T_{\text{я}}(i, j) \right) + b + \varepsilon(i), \quad (4)$$

Также, для удобства вычислений, сумма произведений коэффициентов и предикторов (отклонение радиоярких температур) множественной линейной регрессии подвергается переобозначению, с целью включения в нее константы b :

$$b = x(1), \quad (5)$$

$$\Delta T_{\text{я}}(i, 1) = 1, \quad (6)$$

$$\Delta W(i) = \left(\sum_{j=1}^{k+1} x(j) * \Delta T_{\text{я}}(i, j) \right) + \varepsilon(i), \quad (7)$$

Данное выражение можно записать в матричной форме:

$$\overline{\Delta W} = \widehat{\Delta T}_{\text{я}} \vec{x} + \vec{\varepsilon}, \quad (8)$$

Для решения данной системы линейных алгебраических уравнений используется метод наименьших квадратов (МНК). Матричную форму системы можно переписать в следующем виде:

$$\overline{\Delta W}' = \widehat{\Delta T}_{\text{я}} \vec{x}, \quad (9)$$

Вектор случайной ошибки можно выразить в следующем виде:

$$\vec{\varepsilon} = \overline{\Delta W} - \overline{\Delta W}' = \overline{\Delta W} - \widehat{\Delta T}_r \vec{x}, \quad (10)$$

Суть МНК состоит в выборе такого вектора коэффициентов модели \vec{x} , при которых квадрат вектора ошибок модели $\vec{\varepsilon}$ будет минимальным. Это значит, что для квадрата вектора ошибок множественной линейной регрессии $\vec{\varepsilon}$, имеющего вид:

$$RSS = \vec{\varepsilon}^T \vec{\varepsilon} = (\overline{\Delta W} - \widehat{\Delta T}_r \vec{x})^T (\overline{\Delta W} - \widehat{\Delta T}_r \vec{x}), \quad (11)$$

производная должна быть равна нулю.

$$\frac{dRSS}{d\vec{x}} = 0, \quad (12)$$

При подстановке значения RSS в производную и произведении необходимых вычислений, получается выражение:

$$(\widehat{\Delta T}_r^T \widehat{\Delta T}_r) \vec{x} = \widehat{\Delta T}_r^T \overline{\Delta W}, \quad (13)$$

Из этого следует, что коэффициенты множественной линейной регрессии получаются по формуле:

$$\vec{x} = (\widehat{\Delta T}_r^T \widehat{\Delta T}_r)^{-1} \widehat{\Delta T}_r^T \overline{\Delta W}, \quad (14)$$

После вычисления, полученный вектор коэффициентов регрессии используется для подсчетов значений водозапаса облаков для измеренных радиоярких температур:

$$\overline{W} = \widehat{\Delta T}_r \vec{x} + W_{cp}, \quad (15)$$

$$\Delta T_r(i, j) = T_r(i, j) - T_{r, cp}(j), \quad (16)$$

где $W_{\text{ср}}$ и $T_{\text{яср}}(j)$ – средние значения водозапаса облаков и радиоярких температур обучающего массива соответственно.

Таким образом, получение необходимых результатов состоит из двух этапов:

- вычисление коэффициентов множественной линейной регрессии на основе «обучающих» данных;
- определение значений водозапаса облаков из значений ярких температур, используя полученные коэффициенты.

4.4. Статистическая модель атмосферы и множественное моделирование измерений

Для расчета коэффициентов регрессионного соотношения необходимо провести множественное моделирование измерений яркой температуры нисходящего микроволнового излучения в спектральных каналах радиометра. В настоящей работе мы рассматриваем только режим зондирования в зенитном направлении, поэтому о моделировании измерений при различных зенитных углах речь не идет. В качестве исходных данных для моделирования была использована статистическая модель атмосферы.

Статистическая модель представляет собой набор 2311 состояний системы атмосфера – поверхность. Каждое состояние атмосферы включает в себя профили температуры (К), отношения смеси водяного пара, озона, закиси азота, метана (г/г), водности облаков (г/м³) и параметры поверхности (одновременно оба варианта – и для водной поверхности, и для суши): температуру поверхности (водной или суши), коэффициенты разложения по эмпирическим ортогональным функциям излучательной способности поверхности суши в ИК области спектра, скорость приводного ветра и соленость воды. Последние два параметра позволяют рассчитать излучательную способность водной поверхности в МКВ области спектра.

Данная модель была скомпилирована коллективом авторов СПбГУ в ходе исследований, посвященных космическим экспериментам с аппаратурой ИКФС-2 и МТВЗА на российском спутнике Метеор [24, 25]. При компиляции за основу была взята известная статистическая модель состояний атмосферы TIGR [26], а также база данных, созданная на основе судовых аэрологических и гидрометеорологических измерений, измерений на островных станциях, и предоставленная сотрудниками международного Центра по окружающей среде и дистанционному зондированию имени Нансена (Петербургский Нансен-центр). Эта база данных использовалась при разработке и тестировании нейронно-сетевого алгоритма обработки спутниковых измерений радиотеплового излучения [27,28,29].

Входными данными для расчетов являлись профили давления, температуры, содержания водяного пара и влажности облаков из статистической модели, выходными - значения яркостной температуры микроволнового излучения в каналах радиометра RPG-НАТPRO. Результаты этих расчетов, выполненных разработчиком соответствующего программного обеспечения В.С.Косцовым, были предоставлены автору настоящей работы для анализа и дальнейшего использования в виде следующих наборов данных:

- 2311 значений водозапаса облаков
- 2311 значений радиояркостной температуры в 14 частотных каналах (7 каналов в линии поглощения водяного пара и 7 каналов – в полосе поглощения кислорода).

Совокупность этих наборов данных ниже будем обозначать как AL.

5. Исследование погрешностей определения водозапаса алгоритмом МЛР на основе модельных расчетов

Исследование погрешностей определения водозапаса алгоритмом МЛР происходило следующим образом: для построения коэффициентов МЛР по алгоритму, описанному в разделе 4.1., использовался специально выбранный массив данных, в дальнейшем обозначаемый как обучающий массив. Далее, выбирался второй массив данных, обозначаемый в дальнейшем как тестовый массив. Набор радиоярких температур из тестового массива использовался для вычисления значений водозапаса по алгоритму МЛР. Полученные результаты сравнивались с исходными значениями водозапаса из тестового массива.

Для исследования погрешности метода МЛР использовались данные из массива AL. Из массива отсеивались данные для значений водозапаса более $0,4 \text{ кг/м}^2$, так как эти значения водозапаса облаков в большинстве случаев соответствуют периодам дождя, не рассматриваемым в данной работе. Для расчетов методом МЛР, из массива AL для каждой реализации использовались 7 значений радиоярких температур, соответствующих частотным каналам спектральной линии поглощения водяного пара.

В первом варианте расчетов в качестве обучающего и тестового массива использовался массив AL.

Табл. 4. Характеристики массива LREA AL и полученные на его основе коэффициенты МЛР.

Номера значений водозапаса, используемых для обучения	1-2311
Средние значения радиоярких температур обучающего массива	22,24 ГГц: 40,704 К 23,04 ГГц: 40,014 К 23,84 ГГц: 36,371 К 25,44 ГГц: 29,864 К 26,24 ГГц: 27,942 К 27,84 ГГц: 26,055 К 31,40 ГГц: 26,549 К
Коэффициенты МЛР	$x_1 = -0,02472108$ $x_2 = 0,02329485$ $x_3 = 0,02895755$ $x_4 = -0,00265914$ $x_5 = -0,04173094$ $x_6 = -0,06082760$ $x_7 = 0,08038193$ $b = 0,00000107$
Среднее значение водозапаса облаков обучающего массива	0,138 кг/м ²

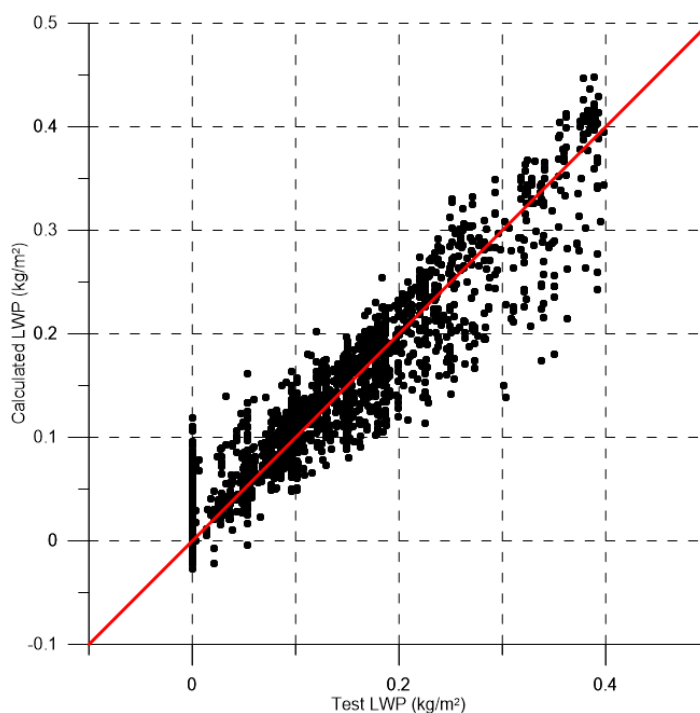


Рис. 7. Сравнение тестовых и рассчитанных значений водозапаса облаков (обучение и тест: AL).

Коэффициенты МЛР и сравнение полученных результатов с значениями водозапаса облаков из тестового массива представлены в Табл. 4 и на Рис.7 соответственно. Данное сравнение показывает значимый разброс рассчитанных значений водозапаса облаков относительно тестовых: с завышением – для значений, близких к 0 кг/м^2 , и с занижением – при больших значениях.

Для анализа отклонений рассчитанного набора значений водозапаса облаков от тестовых значений, рассмотрим их распределения (Рис. 8 и Рис. 9):

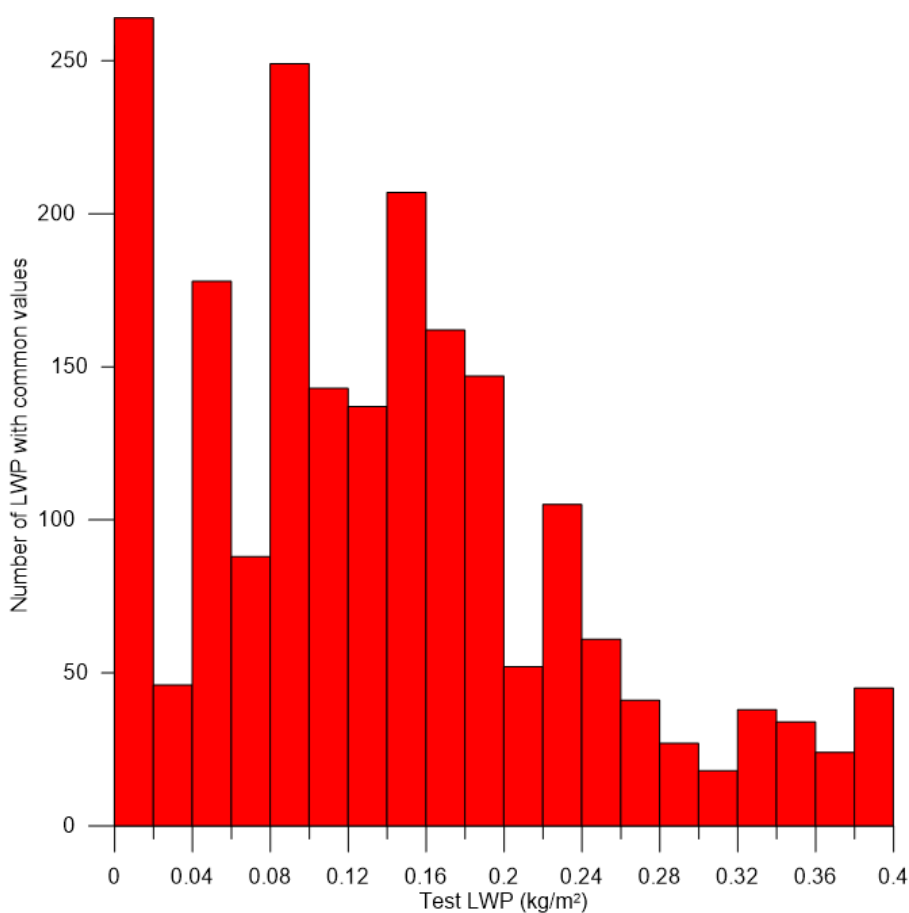


Рис. 8. Распределение тестовых значений водозапаса облаков из массива AL.

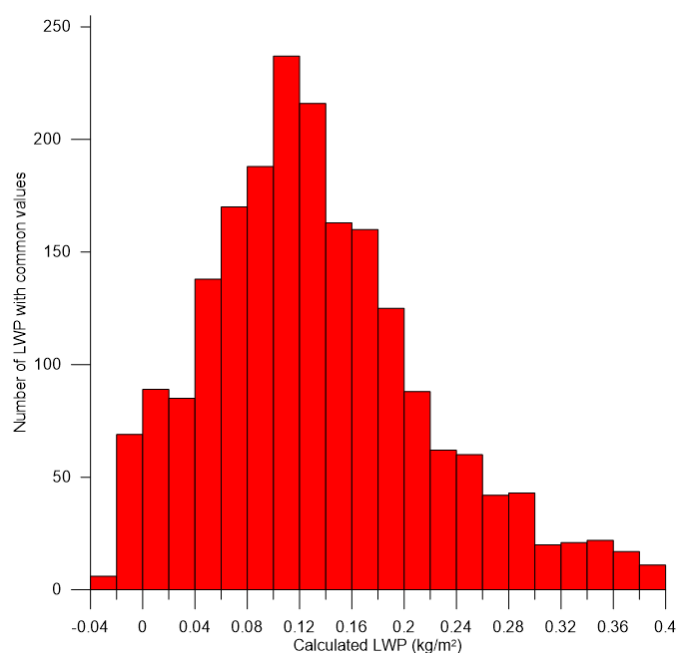


Рис. 9. Распределение значений водозапаса облаков, рассчитанных с помощью МЛР (обучающий массив AL).

Рис. 8 и Рис. 9 наглядно демонстрируют отличия в распределениях двух сравниваемых массивов. Распределение массива тестовых данных показывает наличие визуально определяемых выбросов по всей его структуре. Распределение данных, построенных с помощью множественной линейной регрессии, напротив, практически не имеет выбросов в своей структуре. Статистические характеристики сравнения распределений указаны в Табл. 5:

Табл. 5. Статистические параметры расхождения наборов значений водозапаса из массива LREA AL и значений, построенных методом МЛР.

Среднее расхождение между значениями водозапаса облаков, рассчитанными МЛР, и тестовыми значениями	$3,4 * 10^{-11}$ кг/м ²
Среднеквадратическое отклонение расхождение между значениями водозапаса облаков, рассчитанными МЛР, и тестовыми значениями	0,033 кг/м ²
Коэффициент корреляции значений водозапаса облаков, рассчитанных МЛР, с тестовыми значениями	$0,940 \pm 0,008$

Как можно заметить среднее расхождение между двумя наборами значений пренебрежимо мало, что ожидаемо при использовании одного и

того же набора данных для обучения и теста алгоритма МЛР. Однако, значение среднеквадратического расхождения является большим: 24% от средней величины водозапаса по всему ансамблю. Тем не менее, несмотря на наличие больших вариаций в расхождении значений двух наборов, коэффициент корреляции показывает их сильную зависимость.

Во втором варианте расчетов были сформированы два обучающих массива, параметры которых представлены в Табл. 6:

Табл. 6. Характеристики массивов LREA M1 и LREA M2, а также полученных на их основе коэффициенты МЛР.

Название обучающего массива	M1	M2
Номера значений водозапаса, используемых для обучения	1-700	701-1400
Средние значения радиоярких температур обучающего массива	22,24 ГГц: 68,324 К 23,04 ГГц: 66,398 К 23,84 ГГц: 58,640 К 25,44 ГГц: 44,576 К 26,24 ГГц: 40,139 К 27,84 ГГц: 35,051 К 31,40 ГГц: 32,750 К	22,24 ГГц: 39,976 К 23,04 ГГц: 39,336 К 23,84 ГГц: 35,717 К 25,44 ГГц: 29,197 К 26,24 ГГц: 27,260 К 27,84 ГГц: 25,354 К 31,40 ГГц: 25,830 К
Коэффициенты МЛР	$x_1 = -0,01295935$ $x_2 = 0,0109186$ $x_3 = 0,0219568$ $x_4 = -0,00994763$ $x_5 = -0,0396028$ $x_6 = -0,03059489$ $x_7 = 0,0682315$ $b = -0,00000109$	$x_1 = -0,02434013$ $x_2 = 0,02565088$ $x_3 = 0,02362996$ $x_4 = 0,00753212$ $x_5 = -0,0521208$ $x_6 = -0,06084895$ $x_7 = 0,08350834$ $b = -0,00000027$
Среднее значение водозапаса облаков обучающего массива	0,121 кг/м ²	0,143 кг/м ²

Табл. 6 показывает, что наборы данных M1 и M2 имеют существенные различия: средние радиояркие температуры массива M1 на 23-41% больше средних радиоярких температур массива M2, однако среднее

значение водозаса облаков набора M1 меньше на 15%. Подобное различие делает данные массивы весьма удобными вариантами для изучения зависимости МЛР от «обучающего» массива.

Для теста МЛР по данным двум наборам использовались значения радиояркой температуры и водозаса облаков из массива AL с номерами от 1401 до 2100 (массив этих значений обозначен как TE). Результаты сравнения вышеуказанных значений водозаса с соответствующими им значениями, рассчитанными по методу МЛР, представлены на Рис. 10 и Рис.11:

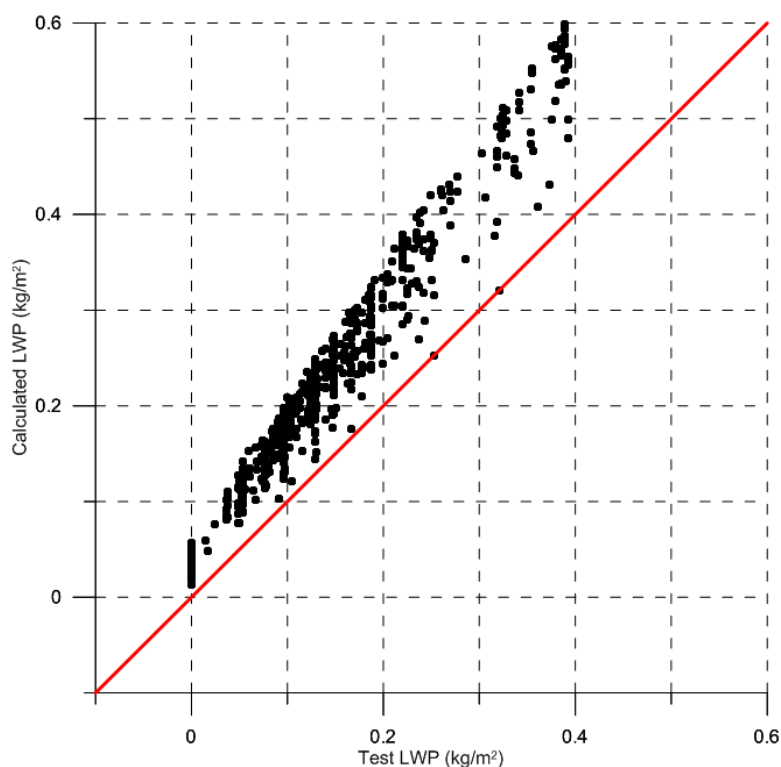


Рис. 10. Сравнение тестовых и рассчитанных значений водозаса облаков (обучающий массив: M1).

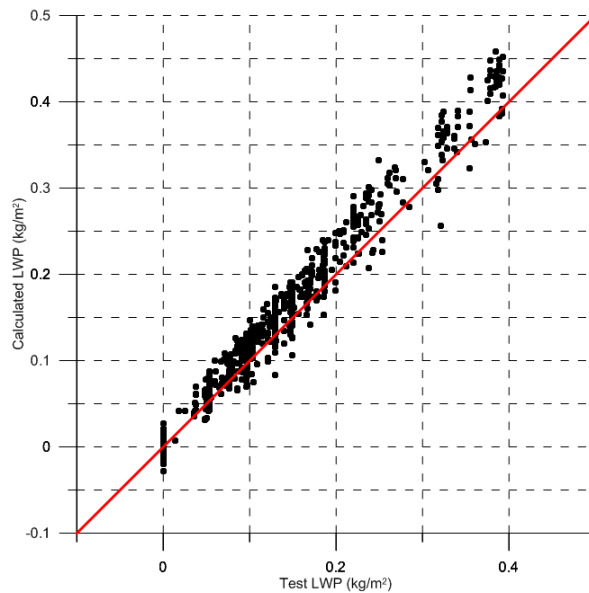


Рис. 11. Сравнение тестовых и рассчитанных значений водозапаса облаков (обучающий массив: M2).

На рисунках заметно завышение значений водозапаса облаков, рассчитанных методом МЛР, относительно тестовых значений, причем данное завышение сильно заметно при обучении МЛР массивом M1 и слабо выражено в случае обучения массивом M2. Ниже, на Рис. 12 и Рис. 13, представлены распределения тестовых и рассчитанных значений водозапаса в случаях использования массивов M1 и M2 для обучения:

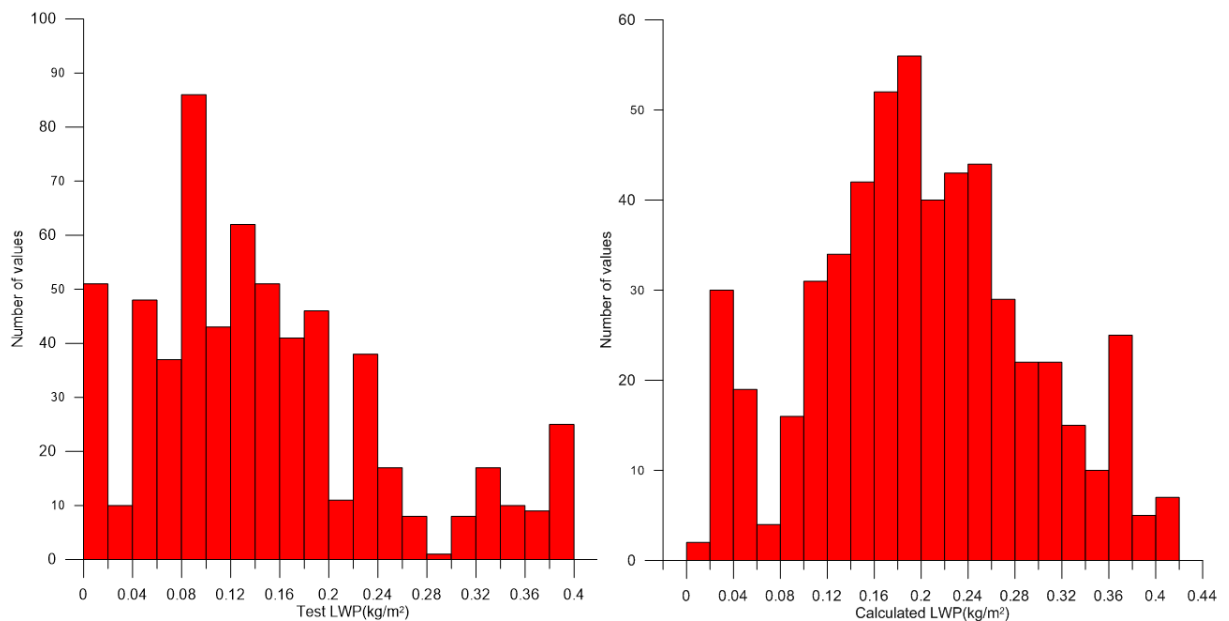


Рис. 12. Распределение тестовых значений водозапаса облаков и значений, полученных при расчете методом МЛР (обучающий массив: M1).

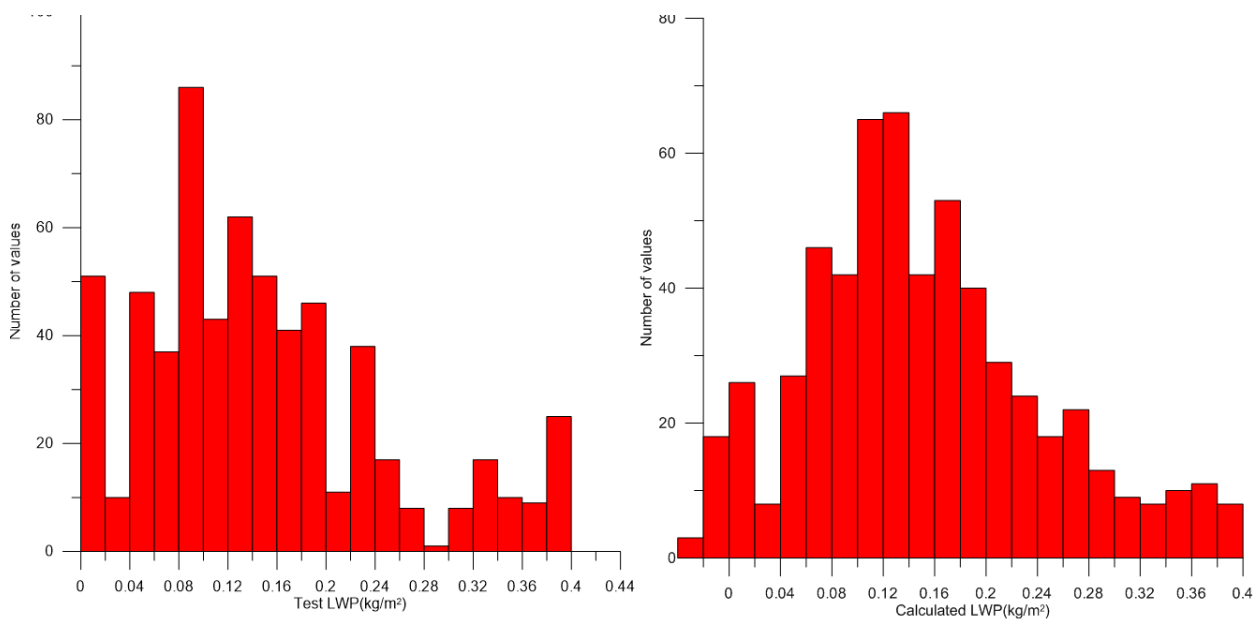


Рис. 13. Распределение тестовых значений водозапаса облаков и значений, полученных при расчете методом МЛР (обучающий массив: M2).

Распределения рассчитанных и тестовых данных для обоих случаев (Рис.12-13) показывают, что в тестовом распределении ТЕ также присутствуют выбросы, но они значительно меньше, чем в случае использования массива AL. Из рис.12-13, можно заметить, что большее сходство структуры распределений тестового и рассчитанного методом МЛР наборов демонстрируется при обучающем массиве M2. Статистические характеристики расхождений тестовых и рассчитанных методом МЛР значений для обоих обучающих массивов приведены в Табл.7:

Табл. 7. Статистические параметры расхождения распределений наборов значений водозапаса из массива LREA ТЕ, а также значений, построенных методом МЛР.

Обучающий массив	M1	M2
Среднее расхождение между значениями водозапаса облаков, рассчитанными МЛР, и тестовыми значениями	0,089 кг/м ²	0,017 кг/м ²
Среднеквадратическое отклонение расхождения между значениями водозапаса облаков, рассчитанными МЛР, и тестовыми значениями	0,041 кг/м ²	0,020 кг/м ²
Коэффициент корреляции значений водозапаса облаков, рассчитанных МЛР, с тестовыми значениями	0,980±0,008	0,985±0,007

Данные параметры показывают, что большее среднее расхождение, а также большее среднеквадратическое расхождение получены при обучающем массиве M1. Среднее расхождение результатов, полученных при расчете МЛР, обученной массивом M1, с тестовыми данными в 5,2 раза больше, чем среднее расхождение результатов, полученных при расчете МЛР, обученной массивом M2, с тестовыми данными. Корреляция для данного массива также показывает меньшую взаимосвязь тестовых и рассчитанных наборов значений из обоих случаев. Таким образом, проведенные тестовые расчеты показывают, что:

- 1) использование различных обучающих массивов может приводить к существенно разным значениям погрешности МЛР,
- 2) при обработке экспериментальных данных на первом этапе исследований, цель которого – анализ точности МЛР, представляется целесообразным использовать несколько наборов регрессионных коэффициентов.

6. Применение множественной линейной регрессии к данным, полученным МКВ радиометром RPG НАТPRO.

6.1. Исследование результатов, полученных МЛР в периоды с большой вариацией систематической составляющей погрешности определения водозапаса облаков

Как уже было сказано в предыдущем разделе для обучения были использованы массивы M1 и M2. Коэффициенты и средние значения параметров двух наборов МЛР были использованы для определения значений водозапаса облаков по измерениям радиояркостных температур с помощью МКВ радиометра RPG НАТPRO. Для удобства, в дальнейшем МЛР будут обозначены как LREA M1 и LREA M2. Результаты данных МЛР также сравнивались с квадратической регрессией (QREA) – алгоритмом, лежащим в

основе программного обеспечения (ПО) поставляемого с аппаратурой НАТPRO. При анализе данных, полученных прибором, был выбран период с большой вариацией систематической составляющей погрешности водозапаса облаков – 18 августа – 7 сентября 2013 года. Данная вариация представлена на Рис.4-5. На данном рисунке видно, что изменение систематической составляющей на данный период превысило 100 г/м^2 . Для анализа были выбраны данные прибора с пятиминутным шагом измерений. Из вышеуказанного периода анализировались непосредственно моменты вариаций систематической составляющей: 09:00-17:00 19 августа 2013 года и 13:00-15:00 4 сентября 2013 года. Данные периоды представлены на Рис.14-15:

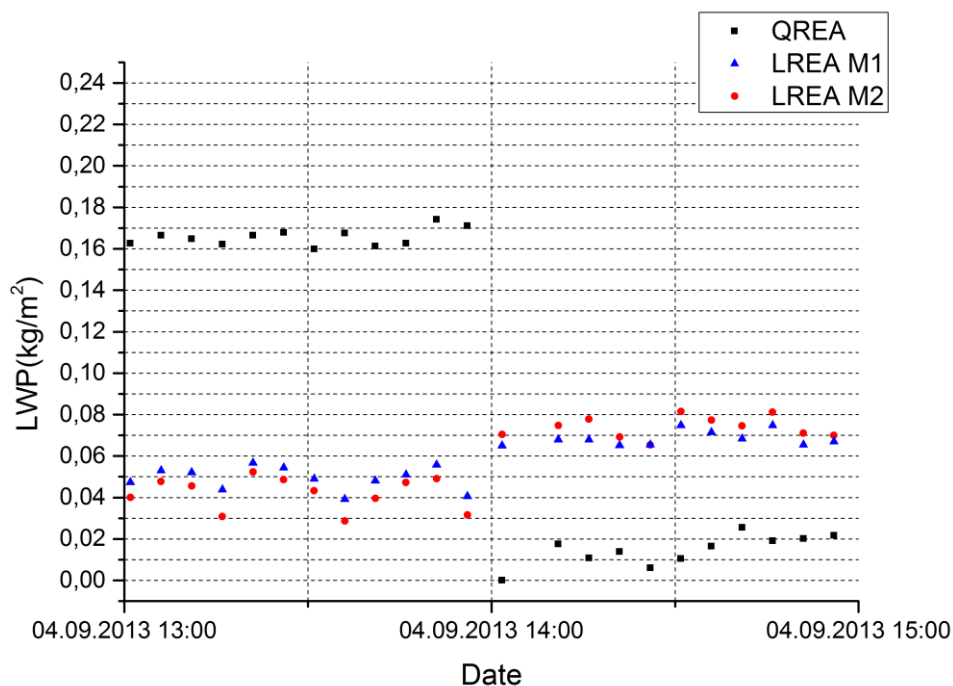


Рис. 14. Ход водозапаса для 13:00-15:00 4 сентября 2013 года для значений, рассчитанных методом квадратической регрессии (QREA), а также с помощью LREA M1 и LREA M2.

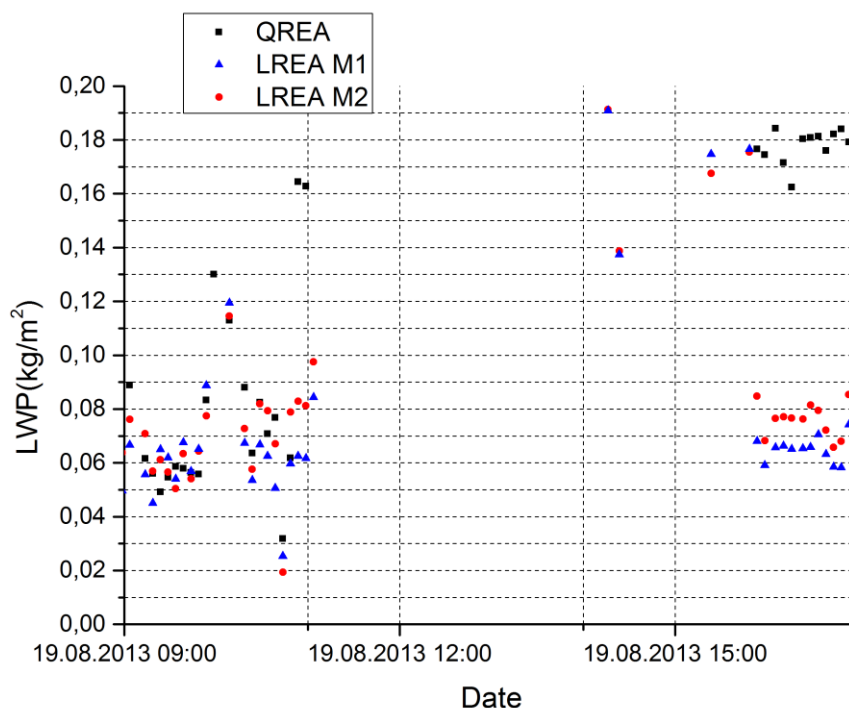


Рис. 15. Ход водозапаса для 09:00-17:00 19 августа 2013 года для значений, рассчитанных методом квадратической регрессии (QREA), а также с помощью LREA M1 и LREA M2.

Из Рис. 14-15 видно, что вариация систематической составляющей погрешности определения водозапаса облаков в рассматриваемых периодах времени для LREA M1 и LREA M2 существенно меньше, чем для QREA. Необходимо привести оценку данной вариации:

Табл. 8. Вариация систематической составляющей погрешности определения водозапаса облаков в период 13:00-15:00 4 сентября 2013 года.

Алгоритм	QREA	LREA M1	LREA M2
Значение вариации систематической составляющей погрешности определения водозапаса облаков	-0,150 кг/м ²	0,019 кг/м ²	0,032 кг/м ²

В Табл.8 представлена оценка вариации систематической составляющей погрешности в период 13:00-15:00 4 сентября 2013. График хода водозапаса облаков (Рис.14) показывает, ввиду отсутствия облачности, хорошо заметную вариацию систематической составляющей погрешности. Также заметно (из Табл. 8 и Рис. 14), что характер вариации систематической составляющей погрешности у квадратической и линейной регрессий разный: в результате

вариации у QREA произошло уменьшение значений водозапаса облаков, а в случаях LREA M1 и LREA M2 – увеличение значений. Величины вариаций систематической составляющей погрешности расчета водозапаса облаков методом МЛР LREA M1 и LREA M2 значительно меньше данной величины в методе QREA: на 87% и 79% соответственно. Также, величина вариации в методе LREA M2 больше от величины вариации в методе LREA M1 в 1,68 раз. Таким образом, в данный период наименьшая вариация систематической составляющей погрешности определения водозапаса облаков в период 13:00-15:00 4 сентября 2013 наблюдается при использовании LREA M1.

Табл. 9. Вариация систематической составляющей погрешности определения водозапаса облаков в период 09:00-17:00 19 августа 2013 года.

Алгоритм	QREA	LREA M1	LREA M2
Значение вариации систематической составляющей погрешности определения водозапаса облаков	0,088 кг/м ²	-0,006 кг/м ²	-0,026 кг/м ²

В период 09:00-17:00 19 августа 2013 года определить величину вариации систематической составляющей погрешности определения водозапаса облаков оказалось значительно труднее, по причине прохождения облачного массива над прибором в рассматриваемый промежуток времени. Однако, ориентируясь на период времени, рассматриваемый выше, удалив из анализа значения водозапаса для периодов измерений с большой облачностью, данные величины были определены и представлены в Табл. 9. Исходя из Табл. 9, характер вариации систематической составляющей погрешности определения водозапаса облаков в период 09:00-17:00 19 августа 2013 года следующий: Для QREA значения водозапаса облаков увеличились, а для LREA M1 и LREA M2 – уменьшились. Величины вариаций для обоих методов МЛР отличаются от величины вариации для QREA следующим образом: величина вариации для LREA M1 меньше на 97%, а величина вариации для LREA M2 меньше на 70%. Значение вариации систематической составляющей погрешности определения водозапаса

облаков для метода LREA M2 больше аналогичного значения для метода LREA M1 в 4,3 раза. Следовательно, и в период 09:00-17:00 19 августа 2013 года наименьшая вариация систематической составляющей погрешности определения водозапаса облаков достигается при использовании LREA M1.

Основной вывод из полученных результатов: отклик МЛР на сбои во входных данных на порядок меньше, чем отклик квадратической регрессии.

6.2. Сопоставление результатов МЛР с результатами, полученными другими алгоритмами, для облачности с большим и малым водозапасом облаков

В данном разделе работы рассматривается период 15 сентября – 15 октября 2014 года. Для анализа хода водозапаса облаков данный период используются измерения радиояркостной температуры с шагом в 2 минуты. Для построения хода водозапаса по алгоритмам LREA M1, LREA M2 и QREA, необходимо определить систематическую составляющую погрешности водозапаса облаков, и убрать ее из рассматриваемых результатов расчетов водозапаса облаков по каждому методу. Для определения систематической составляющей необходимо рассмотреть близкий к изучаемому интервалу безоблачный период времени. Подобный безоблачный период присутствует в самом рассматриваемом промежутке: это период времени с 12 часов 15 сентября по 12 часов 18 сентября 2014 года. Данный период представлен на Рис. 16:

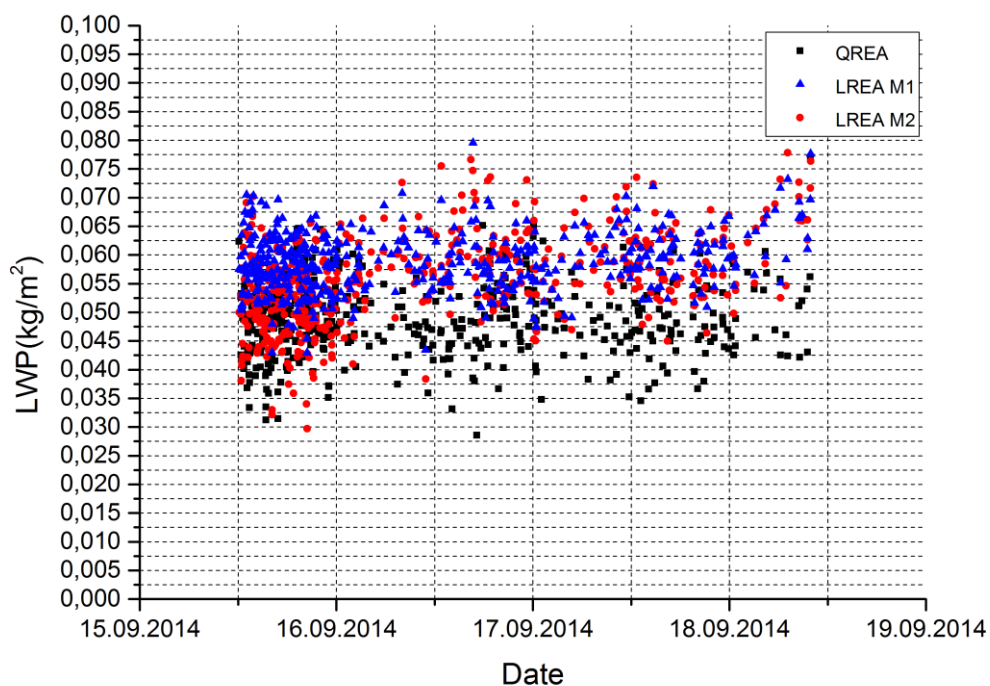


Рис.16. Ход водозапаса облаков для методов LREA M1, LREA M2 и QREA в период 12:00 15.09.14 – 12:00 18.09.14

Рис. 16 показывает, что систематическая составляющая погрешности составляет десятки г/м^2 . Более точно ее величину можно определить, подсчитав среднее значение водозапаса по всему периоду для каждого из алгоритмов:

Табл.10. Систематическая составляющая погрешности определения водозапаса облаков для LREA M1, LREA M2 и QREA.

Алгоритм	QREA	LREA M1	LREA M2
Систематическая составляющая погрешности определения водозапаса облаков	0,049 кг/м^2	0,059 кг/м^2	0,056 кг/м^2
Погрешность систематической составляющей погрешности определения водозапаса облаков	0,0003 кг/м^2	0,0004 кг/м^2	0,0002 кг/м^2

Согласно Табл. 10, наименьшее значение систематической составляющей погрешности определения водозапаса облаков в рассматриваемый период имеет место в методе QREA. Однако, исходя из результатов предыдущего раздела и Рис. 4-5, известно, что данный метод может давать значительную вариацию систематической составляющей, имеющей, ко всему прочему, случайный характер. Поэтому, если оценивать значение систематической составляющей погрешности определения водозапаса облаков среди двух МЛР, то наименьшее значение систематической составляющей имеет метод LREA M2.

Получив систематическую составляющую погрешности определения водозапаса облаков для методов LREA M1, LREA M2 и QREA, и исключив ее из результатов расчетов водозапаса облаков по каждому методу, был построен сравнительный график хода водозапаса облаков для алгоритмов LREA M1, LREA M2, QREA, а также РНА – физического алгоритма определения водозапаса облаков. Результаты, полученные физическим алгоритмом, были предоставлены автору настоящей работы В.С.Косцовым. Каждый из регрессионных алгоритмов сравнивался с физическим алгоритмом, выступающим в качестве «эталона» для анализа. Весь ход водозапаса облаков в период 15 сентября – 15 октября 2014 года представлен на Рис.17:

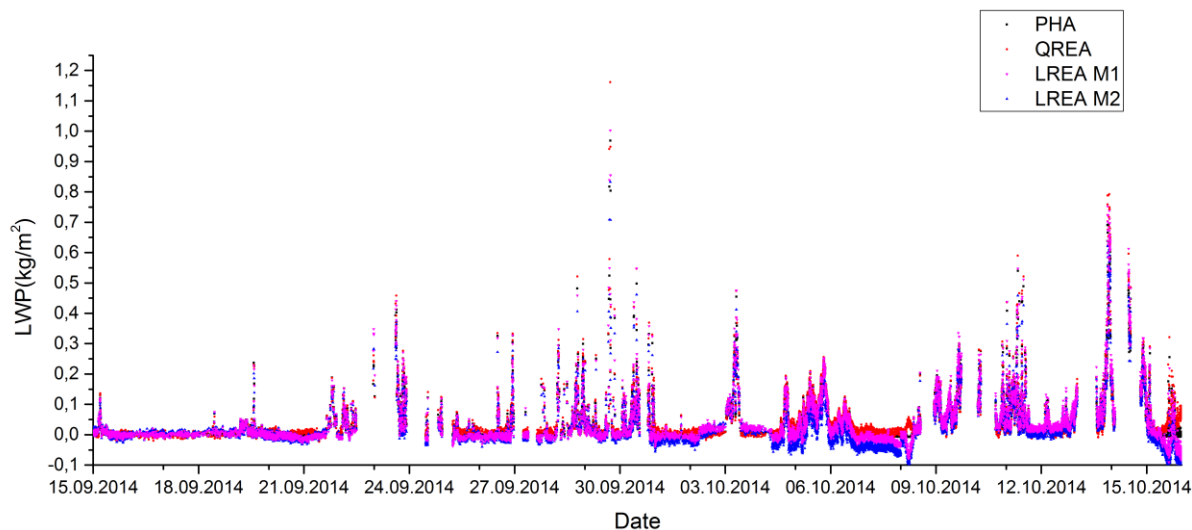


Рис.17 Ход водозапаса облаков в период 15 сентября – 15 октября 2014 года для 4 методов.

На данном рисунке можно заметить периоды, когда значения водозапаса, полученные МЛР, значительно меньше, а значения, полученные квадратической регрессией больше значений, полученных физическим алгоритмом. Данные периоды времени характеризуют сбои в работе прибора. Один из таких сбоев представлен на Рис.18:

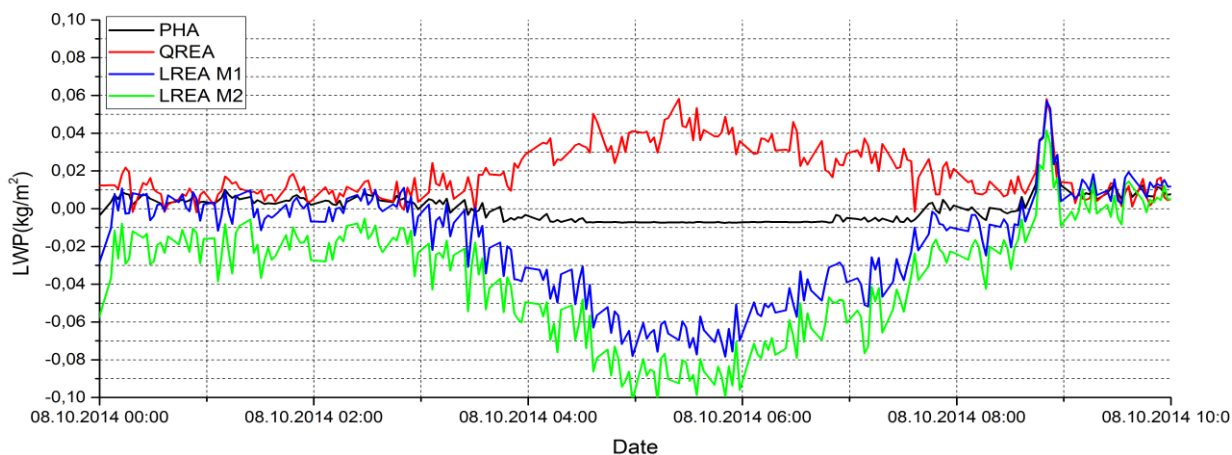


Рис.18. Сбой в работе МКВ радиометра RPG НАТPRO.

Для анализа работы методов в данном периоде времени, необходимо рассмотреть в данном периоде отдельные промежутки, характеризующиеся различными значениями водозапаса облаков. По этой причине для анализа были выбраны следующие промежутки времени: 12:00 12.10.2014 – 00:30

13.10.2014 – с малым водозapasом; 13:41 13.10.2014 – 00:00 14.10.2014 – с большим водозapasом. Данные периоды представлены на Рис.19 и Рис.20.

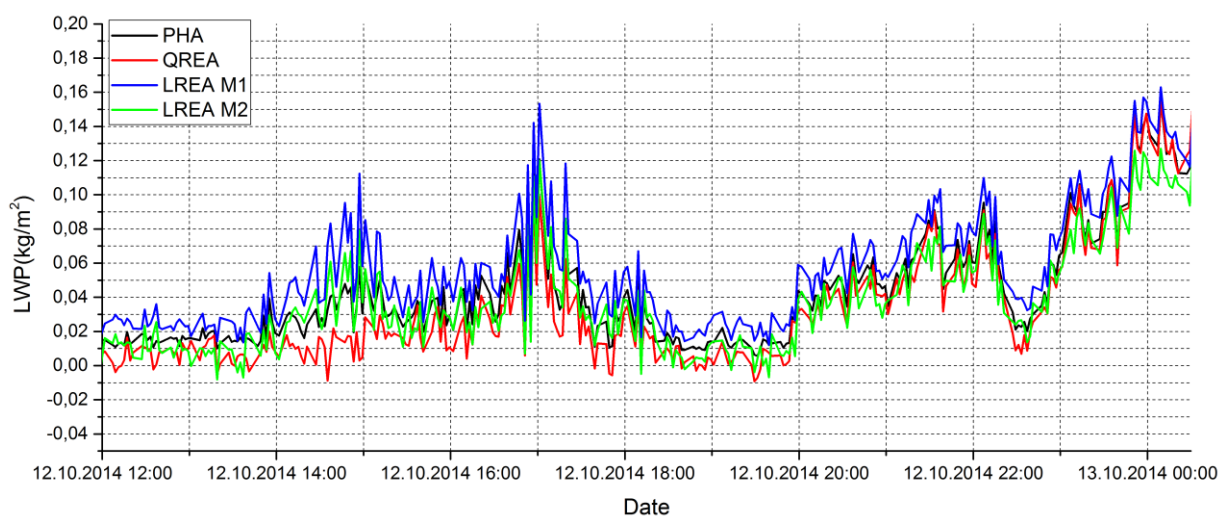


Рис.19. Ход водозapasа облаков для 4 методов в период 12:00 12.10.2014 – 00:30 13.10.2014.

Табл.11. Анализ расхождения значений водозapasа облаков, рассчитанных 4 методами, в период 12:00 12.10.2014 – 00:30 13.10.2014.

Расхождение	QREA-PHA	LREA M1-PHA	LREA M2-PHA	LREA M2-LREA M1
Среднее расхождение между рассматриваемыми алгоритмами	-0,010 кг/м ²	0,013 кг/м ²	-0,006 кг/м ²	-0,019 кг/м ²
СКВ расхождение между рассматриваемыми алгоритмами	0,008 кг/м ²	0,008 кг/м ²	0,009 кг/м ²	0,005 кг/м ²
Коэффициент корреляции между рассматриваемыми алгоритмами	0,982±0,011	0,973±0,013	0,965±0,015	0,993±0,06

Из Рис.19 и Табл.11 видно, что для малооблачного периода измерений самое минимальное расхождение с физическим методом у метода LREA M2. У LREA M1 аналогичное расхождение соразмерно с расхождением QREA-PHA. Все регрессионные методы имеют сильную корреляцию с физическим. Обе МЛР относительно сильно различаются между собой в среднем - на 0,019 кг/м², однако имеют высокую взаимную корреляцию. Все

вышеперечисленное говорит о том, что для расчета водозапаса облаков в рассмотренный период предпочтительно использовать LREA M2.

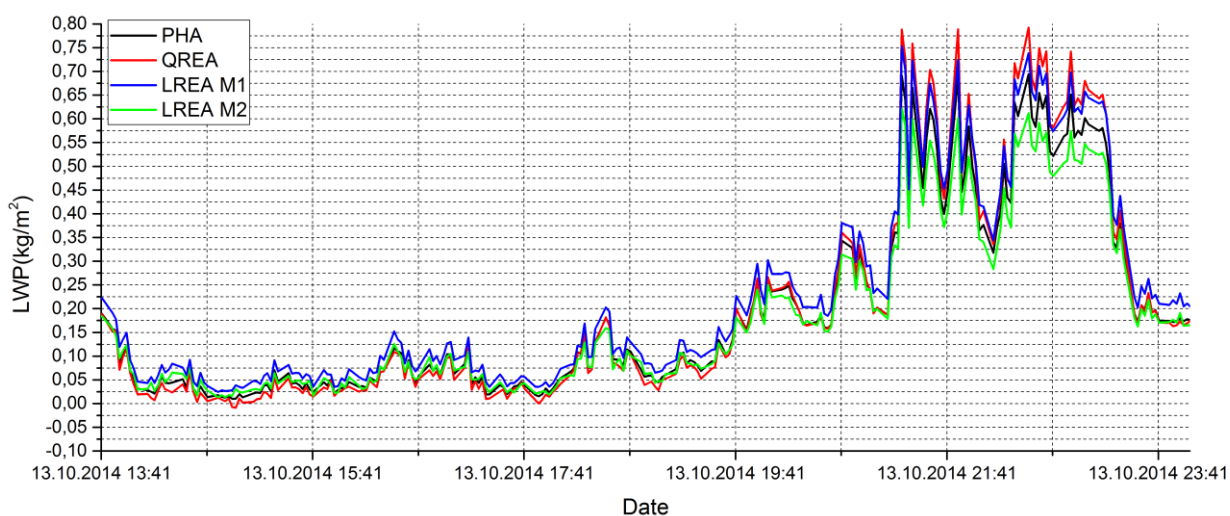


Рис.20. Ход водозапаса облаков для 4 методов в период 13:41 13.10.2014 – 00:00 14.10.2014.

Табл.12. Анализ расхождения значений водозапаса облаков, рассчитанных 4 методами, в период 13:41 13.10.2014 – 00:00 14.10.2014.

Расхождение	QREA-PHA	LREA M1- PHA	LREA M2- PHA	LREA M2-LREA M1
Среднее расхождение между рассматриваемыми алгоритмами	0,007 кг/м ²	0,030 кг/м ²	-0,012 кг/м ²	-0,042 кг/м ²
СКВ расхождение между рассматриваемыми алгоритмами	0,030 кг/м ²	0,013 кг/м ²	0,023 кг/м ²	0,033 кг/м ²
Корреляция между рассматриваемыми алгоритмами	0,999±0,002	0,999±0,002	0,999±0,002	1,000±0,001

Рис.20 и Табл.12 представляют данные для периода с мощной облачностью. В данном случае все регрессионные методы имеют одинаково

сильную корреляцию с физическим методом. Данная корреляция сильнее чем в предыдущем случае. Однако наименьшее расхождение в данном промежутке наблюдается между методами QREA и РНА. Тем не менее, из методов МЛР наименьшее расхождение с физическим методом имеет LREA M2. Между собой МЛР различаются гораздо сильнее предыдущего случая – на $0,042 \text{ кг/м}^2$, однако их взаимная корреляция тоже сильнее: в данном случае она достигает значения абсолютной корреляции. В итоге, принимая во внимание то, что большая величина вариации систематической составляющей погрешности определения водозапаса облаков методом QREA делает анализ водозапаса облаков затруднительным, можно говорить о том, что для расчета водозапаса облаков в рассмотренный период предпочтительно использовать LREA M2.

7. Основные результаты и выводы

В настоящей диссертационной работе решалась задача реализации метода линейной регрессии для определения водозапаса облаков по наземным измерениям микроволнового излучения аппаратурой НАТPRO, проводились оценки точности этого метода, и результаты определения водозапаса сравнивались с результатами, полученными независимыми методами.

Были получены следующие результаты:

- 1) Погрешность определения водозапаса облаков методом линейной регрессии существенно зависит от качества и полноты «обучающего» ансамбля данных. Численные эксперименты, проведенные на двух обучающих и одном тестовом ансамбле данных показали, что различие в средней погрешности может достигать $0,072 \text{ кг/м}^2$, а в среднеквадратической – $0,019 \text{ кг/м}^2$.
- 2) Обучение на различных по качеству и полноте ансамблях данных влияет и на величину систематической погрешности определения

водозапаса. Согласно сделанным оценкам систематическая составляющая погрешности может изменяться на $0,003 \text{ кг/м}^2$.

- 3) Метод линейной регрессии обладает заметно меньшей чувствительностью к сбоям во входных данных, чем метод квадратической регрессии. Например, в период 4 Сентября 2014 года в результате сбоя в результатах измерений систематическая погрешность метода квадратической регрессии изменилась на $0,15 \text{ кг/м}^2$, а метода линейной регрессии – на $0,019 \text{ кг/м}^2$ и $0,032 \text{ кг/м}^2$ для LREA M1 и LREA M2 соответственно.
- 4) Сравнение обоих регрессионных методов с физическим методом для периода с большим значением водозапаса облаков показало, что среднее расхождение и скв расхождение QREA с методом РНА составляет $0,007 \text{ кг/м}^2$ и $0,030 \text{ кг/м}^2$, тогда как расхождения LREA M2 с методом РНА составляет $-0,012 \text{ кг/м}^2$ и $0,023 \text{ кг/м}^2$. В случае периода с малым значением водозапаса значения следующие: среднее и скв расхождения QREA с РНА составляют $-0,010 \text{ кг/м}^2$ и $0,008 \text{ кг/м}^2$, тогда как расхождение LREA M2 с РНА составляет $-0,006 \text{ кг/м}^2$ и $0,009 \text{ кг/м}^2$.

Принимая все вышеперечисленные результаты во внимание, можно говорить о том, что для расчета водозапаса облаков возможно использовать МЛР в качестве замены множественной квадратической регрессии. Более того, в ходе данной работы, некоторые результаты расчета водозапаса облаков методом МЛР оказались лучше, чем результаты расчета множественной квадратической регрессией. Тем не менее, переход на использование метода МЛР предполагает его доработку с использованием в качестве обучающего массива более качественного и представительного набора данных. Необходимо отметить, что формирование более качественного и представительного обучающего массива выходило за рамки данной работы.

Список литературы:

1. Облака и облачная атмосфера. Справочник. Под редакцией д-ра физ.-мат. наук, профессора И. П. Мазина и д-ра геогр. наук, профессора А. Х. Хгриана. Ленинград Гидрометеиздат 1989 г.
2. Захарова П.В. Лидарно-радиометрические методы контроля водности и водозапаса жидкокапельных облаков. Москва, 1995. С.3
3. Тимофеев Ю.М., Васильев А.В. Теоретические основы атмосферной оптики – СПб.: Наука, 2003 – 474 с., ил.120.
4. Е. Н. Кадыгров, А. Г. Горелик, Т. А. Точилкина Результаты исследования водозапаса облаков радиометрическим комплексом «Микрорадком»// Оптика атмосферы и океана. 27, №6(2014).
5. Караваев Д. М. СВЧ-радиометрические исследования влагозапаса атмосферы и водозапаса облаков, 2010 г.
6. Косцов В.С., Поберовский А.В., Осипов С.И., Тимофеев Ю.М. Комплексная методика интерпретации наземных микроволновых спектральных измерений в задаче определения вертикального профиля содержания озона // Оптика атмосферы и океана. 2012.- Т.25 - №4 - С. 354-360.
7. Radiometer Physics GmbH. [Электронный ресурс] / Режим доступа: <http://www.radiometer-physics.de/rpg/html/Home.html> свободный.- Загл. с экрана.
8. Susanne Crewell, Harald Czekala, Ulrich Löhnert, and Clemens Simmer. Microwave Radiometer for Cloud Cartography: A 22-channel ground-based microwave radiometer for atmospheric research. 2001.
9. Radiometrics Profiler Operator's Manual MP-300A
10. DOMENICO CIMINI, TIM J. HEWISON and LORENZ MARTIN. Comparison of brightness temperatures observed from ground-based microwave radiometers during TUC.2005
11. O. Stähli, A. Murk, N. Kämpfer, C. Mätzler and P. Eriksson. Microwave radiometer to retrieve temperature profiles from the surface to the stratopause. 2013
12. <https://climatedataguide.ucar.edu/climate-data/liquid-water-path-lwp-uwisc-climatology>
13. <http://www.remss.com>
14. <http://www.remss.com/missions/ssmi>
15. <http://www.remss.com/missions/windsat>
16. <http://www.remss.com/missions/amsr>

17. <http://www.remss.com/missions/tmi>
18. Turner D.D., Shepard A. Clough, James C. Liljegren, Eugene E. Clothiaux, Karen E. Cady-Pereira, and Krista L. Gaustad. 2007. "Retrieving Liquid Water Path and Precipitable Water Vapor from the Atmospheric Radiation Measurement (ARM) Microwave Radiometers" *IEEE Transactions On Geoscience And Remote Sensing*, 45(11): 3680-3690. doi 10.1109/TGRS.2007.903703
19. Meijgaard, E., S. Crewell. 2005. "Comparison of model predicted liquid water path with ground-based measurements during CLIWA-NET" *Atmospheric Research* 75(3): 201-226. doi 10.1016/j.atmosres.2004.12.006.
20. Rose, T., S. Crewell, U. Löhnert, and C. Simmer. 2005. "A Network Suitable Microwave Radiometer for Operational Monitoring of the Cloudy Atmosphere." *Atmospheric Research* 75 (3): 183-200.
21. Crewell, S., and U. Loehnert (2003): "Accuracy of cloud liquid water path from ground-based microwave radiometry, 2, Sensor accuracy and synergy." *Radio Science*. 38(3): 8042. doi:10.1029/2002RS002634.
22. Maetzler, C., and J. Morland. 2009. "Refined Physical Retrieval of Integrated Water Vapor and Cloud Liquid for Microwave Radiometer Data" *Ieee Transactions on Geoscience and Remote Sensing*. 47(6): 1585-1594. doi 10.1109/TGRS.2008.2006984
23. Kostsov V.S., Ionov D.V., Biryukov E.Yu., Zaitsev N.A. Cross-validation of two liquid water path retrieval algorithms applied to ground-based microwave radiation measurements by RPG-HATPRO instrument // submitted to International Journal of Remote Sensing. 2017
24. "Разработка и тестирование на модельных данных ПМО повитковых радиационных расчетов для обработки информации аппаратуры МТВЗА" Отчет НИИ физики по договору № Д.02.024, С.-Петербург, 2005.
25. Разработка и тестирование ПМО моделирования и обработки данных глобальных измерений аппаратуры МТВЗА в условиях облачной атмосферы. Отчет НИИФ по договору ХД 02.061. С.-Петербург, 2006, 43 с.
26. TIGR, Thermodynamic Initial Guess Retrieval, [Электронный ресурс] / Режим доступа: <http://ara.lmd.polytechnique.fr/htdocs-public/products/TIGR/TIGR.html>. свободный.- Загл. с экрана.
27. Заболотских Е.В. Восстановление влагосодержания атмосферы и приводного ветра с помощью нейронно-сетевых алгоритмов по спутниковым микроволновым

данным, диссертация на соискание ученой степени кандидата физико-математических наук, С.-Петербург, 2001 г.

28. Заболотских Е.В., Митник Л.М., Бобылев Л.П., Йоханнессен О.М. Разработка и валидация алгоритмов восстановления скорости приводного ветра по данным SSM/I с применением Нейронных Сетей и физических ограничений // Исслед. Земли из космоса. 2000. № 2. С. 62.71.
29. Заболотских Е.В., Тимофеев Ю.М., Успенский А.Б., Митник Л.М., Бобылев Л.П., Йоханнессен О.М., Черный И.В. О точности микроволновых спутниковых измерений скорости приводного ветра, влагосодержания атмосферы и водозапаса облаков // Изв. РАН. Физика атмосферы и океана. 2002. Т. 5. № 38. С. 592.600.