

**Федеральное Государственное бюджетное учреждение
«Главная геофизическая обсерватория им. А.И. Воейкова»
(ФГБУ «ГГО»)**

г. Санкт-Петербург

24 августа 2016 г.

Методическое письмо № 25

**Организация проведения пространственного контроля
метеорологической информации в условиях модернизации
наблюдательной сети**

1 Общие положения

В соответствии с Законом РФ от 26 июня 2008 года № 102-ФЗ «Об обеспечении единства измерений» (глава 1, статья 1), измерения в области гидрометеорологии и мониторинга состояния и загрязнения окружающей среды входят в сферу государственного регулирования обеспечения единства измерений. В период модернизации метеорологической сети особое значение для инструментальных измерений имеют обеспечение и поддержание качества данных.

С 1 ноября 2002 года в РФ принята и введена в действие группа стандартов ГОСТ Р ИСО 5725 под общим названием «Точность (правильность и прецизионность) методов и результатов измерений» с целью применения группы международных стандартов ИСО 5725 для повышения качества измерений. ГОСТ Р ИСО 5725 впервые регламентирует систематические и случайные погрешности и их составляющие.

В публикациях ВМО предписывается применять ко всем метеорологическим данным определенные процедуры контроля качества. В условиях перехода на автоматические средства измерения УГМС особое внимание следует уделять контролю качества результатов наблюдений, причем, как оперативной, так и режимной (климатической) информации.

Вся метеорологическая информация, в т.ч. и поступающая от автоматических метеорологических станций (АМС) без персонала, должна быть подвергнута трем уровням контроля: синтаксическому, семантическому и пространственному. Синтаксический и семантический методы контроля позволяют обнаруживать грубые ошибки, обусловленные случайными факторами. Систематические ошибки наблюдений могут быть выявлены только при пространственном контроле, т.е. при сопоставлении в пространстве значений метеорологических характеристик по группе станций.

В рамках решения задач по обеспечению установленного объема и стандарта качества получаемой информации о состоянии и загрязнении окружающей среды регулярное проведение пространственного контроля материалов наблюдений и анализ его результатов является обязательным направлением деятельности УГМС.

Основные положения по проведению и анализу результатов пространственного контроля изложены в «Рекомендациях по анализу результатов пространственного

контроля режимной метеорологической информации, 1993». Данное методическое письмо разъясняет методику подготовки вспомогательных материалов для проведения пространственного контроля применительно к современным условиям модернизации и технического перевооружения метеорологической наблюдательной сети.

2 Методика подготовки вспомогательных материалов

Эффективность анализа результатов пространственного контроля во многом зависит от правильно и своевременно подготовленных вспомогательных материалов (правильный подбор влияющих станций, оптимальный выбор признака контроля, правильно рассчитанные и обновленные месячные «нормы» станций, статистически значимые месячные коэффициенты уравнений регрессии, рассчитанные для отдельных метеорологических величин и атмосферных явлений и т.д.). Требования по подготовке вспомогательных материалов для проведения контроля, изложены в п.2.3. «Рекомендации по анализу результатов пространственного контроля режимной метеорологической информации», 1993 г.

В зависимости от физико-географических условий контролируемой станции и ее расположения по отношению к влияющим должен быть определен *метод* (для некоторых станций-два метода) *контроля*:

– станция контролируется методом интерполяции, т.е. для нее подобраны влияющие станции в необходимом количестве для всех контролируемых характеристик (кроме атмосферных явлений) – признак контроля «1»;

– станция контролируется только по уравнениям связи (регрессии) – признак контроля «2»;

– станция контролируется методом интерполяции и по уравнениям связи (регрессии) – признак контроля «3»;

– станция контролируется только по уравнениям связи (регрессии), но для контроля атмосферного давления подобрана группа из 3-4 влияющих станций – признак контроля «4».

Автоматизированный пространственный контроль материалов наблюдений выполняется с 1972 г, *нормы и коэффициенты уравнений регрессии*, помещенные в паспортные данные (PP01 и PP02), как правило, устарели и не являются эффективными, что затрудняет анализ результатов пространственного контроля. С целью повышения эффективности проведения контроля и улучшения качества материалов наблюдений требуются проверка и обновление вспомогательных материалов.

Все версии паспортов станций должны сохраняться в отдельном архиве с пояснениями, какие изменения были внесены, и в какой период эти паспорта использовались. Ведение такого архива связано с необходимостью производить повторный контроль старых рядов наблюдений, а для этого требуются именно те паспорта, которые использовались в тот период. Перед редактированием какого-либо паспорта имеющаяся версия должна быть скопирована и сохранена в архиве.

Для редактирования паспортов PP01 и PP02 необходимо пройти по директории METEOWIN\PAS\ и в соответствующих файлах с расширением .txt производить необходимые правки. Полное название паспорта содержит также номер УГМС. Например, PP01_13.txt, PP02_13.txt. Расположение всех файлов ПЕРСОНЫ-МИС прописано в файле METEOWIN\PRG\disk.

2.1 Выбор влияющих станций и станций-аналогов (или величин-аналогов)

Если станция контролируется методом интерполяции необходимо проверить актуальность представленных в паспорте РР01 влияющих станций и при необходимости внести изменения. Так как с 1972 года многие станции были перенесены, закрыты или открывались новые наблюдательные подразделения, изменялась во времени репрезентативность пунктов наблюдений, необходимо пересмотреть и перепроверить список влияющих станций для контролируемой (функционируют ли эти станции, не входят ли они теперь в другое УГМС, являются ли они репрезентативными по отношению к контролируемой станции и т.д.).

В качестве *влияющих* выбирают станции, расположенные на расстоянии не более 150 км от контролируемой станции, в сходных физико-географических условиях. Влияющие станции должны равномерно со всех сторон окружать контролируемую. Оптимальное количество влияющих станций считается 5-6.

Если станция контролируется только по уравнениям связи, необходимо проверить правильность выбора станции-аналога или величины-аналога (правило занесения информации в РР01 приводится ниже).

Если станция контролируется и методом интерполяции, и по уравнениям связи, необходимо проверить актуальность выбора влияющих станций и их координатных номеров, правомерность подобранных групп элементов со станциями-аналогами (величинами-аналогами): первая строка в РР01.

1	2	3	4	5	6	7	8	9	10	11
10	5392091	23	1	5441991	5472091	5371981	5342161	5341991	5282001	2

Рисунок 1 – Пример заполнения первой строки паспорта РР01.

На рисунке 1 цифрами обозначена следующая информация:

- 1** – идентификатор вида информации о нормах станций (2 байта) - «**10**»;
- 2** – координатный номер контролируемой станции (7 байт) – «**5392091**»;
- 3** – номер УГМС (2 байта);
- 4** – признак метода контроля (1байт);
- 5 - 10** – координатные номера 6-ти влияющих станций (по 7 байт каждый);
- 11** – признак характеристики маршрута снегосъемок, по которому заносятся нормы (1 байт): 1 – поле, 2 – лес.

При наличии норм по обоим маршрутам следует выписывать нормы по маршруту, наиболее характерному для всего района, включающего и контролируемую станцию, и влияющие. Если для большинства влияющих станций маршруты снегосъемок проводятся по лесному маршруту, то и норма в паспорте контролируемой станции должна быть занесена для лесного маршрута.

2.2 Нормы основных метеорологических величин

После присвоения контролируемой станции признака метода контроля и обновления информации о влияющих станциях или станциях-аналогах (величин-аналогов) переходят ко второму подготовительному этапу – **заполнение массива с нормами**.

В соответствии с рекомендациями ВМО, *норма* метеорологической характеристики представляет собой *среднее многолетнее значение, рассчитанное за последовательные 30-летние периоды*. Нормы служат контрольной точкой, с которой могут сравниваться недавно проведенные или текущие наблюдения, включая обеспечение проведения пространственного контроля, основанного на отклонениях (аномалиях) текущих данных от нормы.

Важно отметить, что нормы, используемые для пространственного контроля, не всегда совпадают с климатическими нормами. Базисные 30-летние периоды обновляются каждые десять лет для оперативной оценки данных, при этом первым годом периода считается год, заканчивающийся на цифру 1, а последним – заканчивающийся на цифру 0. Таким образом, климатические данные последовательных 30-летних периодов включают следующие временные интервалы: с 1 января 1981 г. по 31 декабря 2010 г., с 1 января 1991 г. по 31 декабря 2020 г. и т.д. Такой 30-летний период является базисным и должен использоваться в анализе и сравнении данных наблюдений во всех странах [The Role of Climatological Normals in a Changing Climate (WMO/ TD-№ 1377)].

Нормы для проведения пространственного контроля основных метеорологических характеристик необходимо пересчитывать каждые десять лет или уточнять при необходимости (перенос станции, изменение репрезентативности пункта наблюдений, повлекшее нарушение однородности рядов). Для эффективного проведения пространственного контроля для каждого НП необходимо рассчитать современные месячные (декадные для снежного покрова) нормы основных метеорологических характеристик за базисный 30-летний период с 1 января 1981 г. по 31 декабря 2010 г., для этого:

- 1) по суточным значениям метеорологической характеристики найти средние месячные значения искомой величины за период 30 лет (например, 30 значений средней месячной температуры воздуха за январь, февраль, март и т.д.);
- 2) по полученным средним месячным значениям какой-либо метеорологической характеристики рассчитать средние многолетние месячные значения, которые являются нормой для конкретного НП.

В качестве примера приводятся средние месячные значения температуры подстилающей поверхности и температуры воздуха по станции Санкт-Петербург, на основании которых рассчитаны соответствующие нормы (таблица 1). Если норма по какой-либо метеорологической величине заносится с точностью до десятых, то предварительные расчеты рекомендуется проводить с точностью до сотых, если норма заносится до целых долей, то расчеты проводятся до десятых долей (как показано в таблице 1: температура подстилающей поверхности заносится в паспорт РР01 до целых, соответственно предварительные расчеты велись до десятых долей, норма по температуре воздуха заносится в РР01 с точностью до десятых долей, соответственно предварительные расчеты осуществлялись с точностью до сотых).

Таблица 1

Средние месячные значения температуры подстилающей поверхности и температуры воздуха за январь для станции Санкт-Петербург, период осреднения с 1981 по 2010 гг.

Год	$t_{п}, ^\circ\text{C}$	$t_{в}, ^\circ\text{C}$
1981	-5,6	-5,60
1982	-12,9	-11,41
1983	-3,5	-3,42
1984	-4,7	-4,63
1985	-15,4	-15,22
1986	-7,5	-7,36
1987	-19,5	-19,42
1988	-6,2	-6,09
1989	-2,4	-2,39
1990	-6,7	-6,63
1991	-5,5	-5,49
1992	-3,9	-3,84
1993	-3,3	-3,24
1994	-4,6	-4,55
1995	-5,8	-5,72
1996	-8,6	-8,54
1997	-7,6	-7,56
1998	-3,1	-3,05
1999	-7,2	-7,22
2000	-5,7	-5,65
2001	-3,7	-3,66
2002	-5,5	-5,52
2003	-11,0	-11,03
2004	-8,7	-8,74
2005	-2,6	-2,62
2006	-6,1	-6,06
2007	-2,8	-2,80
2008	-2,6	-2,62
2009	-4,4	-4,44
2010	-13,8	-13,77
σ	$\pm 4,13$	$\pm 4,05$
\bar{t} (среднее многолетнее значение)	-6,7	-6,61
$3 \cdot \sigma$	$\pm 12,40$	$\pm 12,15$
Допустимые пределы [$\bar{t} - 3\sigma$; $\bar{t} + 3\sigma$]	[-19,1; 5,7]	[-18,76; 5,54]
Норма (без учета аномального года)	-6	-6,2

где, $t_{п}$ – среднее месячное значение температуры подстилающей поверхности, $^\circ\text{C}$;

$t_{в}$ – среднее месячное значение температуры воздуха, $^\circ\text{C}$;

\bar{t} – среднее многолетнее значение метеорологической величины;

σ – среднее квадратическое отклонение (в EXCEL функция СТАНДОТКЛОН) средних месячных значений температуры воздуха и температуры подстилающей поверхности, $^\circ\text{C}$.

При расчете норм из выборки следует исключать аномальные годы, т.е. если разность между средним многолетним месячным значением соответствующей метеорологической величины и значением этой же величины за какой-либо конкретный год (в нашем примере это 1987 г.) превышает утроенное значение среднего квадратического отклонения ($\pm 3\sigma$), такой год исключается. Т.е. находятся допустимые пределы $[\bar{x}-3\sigma; \bar{x}+3\sigma]$ для средних месячных значений метеорологических величин. В примере, представленном в таблице 1, значения за 1987 г. не соответствуют допустимому пределу и из расчета норм должны быть исключены. Несоблюдение этого правила может привести к искажению нормы за счет аномально-холодного или аномально-теплого, или аномально-влажного и т.д. года, что впоследствии негативно повлияет на результаты пространственного контроля.

Нормы, рассчитанные по неполным комплектам данных, вследствие пропусков наблюдений в базисном периоде, также могут быть искаженными. Общее количество полных лет в базисном периоде должно составлять не менее (\geq) 24. При этом обязательно должно выполняться условие, что пропуски не превышают (\geq) 3 последовательных лет подряд. Если в базисном периоде имеется продолжительный период (превышающий допуск) отсутствующих данных, рекомендуется рассчитывать нормы за период, не совпадающий с базисным, но который по полноте данных соответствует требованиям (например, с 1985 по 2015 гг.).

Для короткорядных станций, не имеющих 30-летнего ряда наблюдений, вводится понятие *промежуточной нормы*. В этом случае допускается расчет норм для пространственного контроля за более короткий срок (не менее 10 лет непрерывных наблюдений). При этом по мере увеличения длины ряда наблюдений и соответствующего накопления данных необходимо ежегодно уточнять промежуточные нормы. Для короткорядных станций при анализе результатов пространственного контроля следует учитывать факт использования промежуточной нормы, которая может отразиться на значениях невязок.

В качестве примера заполнения массива с нормами по каждому месяцу в паспорте РР01 «нормы станций», приведена станция Амазар с координатным номером 5392091 (таблица 2).

Таблица 2

Паспорт «нормы станций» для станции Амазар (РР01)

№строки	10	5392091	23	1	5441991	5472091	5371981	5342161	5341991	5282001	2						
1	5392091	1	-298	-31	78	2	5	10	34	7	4	15	23	8	12		
2	5392091	2	-235	-25	70	4	8	12	33	4	5	13	19	19	29	18	29
3	5392091	3	-134	-13	63	12	16	14	41	8	7	17	27	13	20	11	19
4	5392091	4	-11	1	55	33	31	20	56	20	19						
5	5392091	5	79	10	54	64	54	19	65	34	43						
6	5392091	6	150	19	64	84	106	14	66	40	72						
7	5392091	7	188	23	73	77	149	13	66	42	109						
8	5392091	8	155	19	76	59	128	11	65	46	98						
9	5392091	9	81	9	71	42	74	14	63	40	51						
10	5392091	10	-31	-3	69	22	34	12	50	19	15			5			
11	5392091	11	-187	-20	77	5	12	10	42	9	14	3		7	9		4
12	5392091	12	-280	-28	81	2	6	10	38	6	7			13	19		
13	5392091	10	41	83	11	36	85										

Описание заполнения первой строки приведено в п.2.1 данного письма.

В следующие 12 строк заносятся нормы за 12 месяцев (2-13 строки). Каждая строка состоит из 17 столбцов:

- 1) координатный номер контролируемой станции;
- 2) номер месяца;
- 3) норма по температуре воздуха, заносится в градусах с точностью до десятых долей со знаком;
- 4) норма по температуре подстилающей поверхности, заносится в целых градусах со знаком;
- 5) норма по относительной влажности, заносится в процентах с точностью до целых;
- 6) норма по дефициту насыщения водяного пара, заносится в гПа с точностью до десятых;
- 7) норма по парциальному давлению водяного пара, заносится в гПа с точностью до десятых долей;
- 8) норма по скорости ветра, заносится в м/с с точностью до десятых долей;
- 9) норма по количеству общей облачности, заносится в баллах с точностью до десятых долей;
- 10) норма по количеству нижней облачности, заносится в баллах с точностью до десятых долей;
- 11) норма по количеству осадков, заносится в мм с точностью до целых;
- 12) норма по высоте снежного покрова на последний день первой декады, заносится в см с точностью до целых;
- 13) норма по запасу воды в снежном покрове на последний день первой декады, заносится в мм с точностью до целых;
- 14) норма по высоте снежного покрова на последний день второй декады, заносится в см с точностью до целых;
- 15) норма по запасу воды в снежном покрове на последний день второй декады, заносится в мм с точностью до целых;
- 16) норма по высоте снежного покрова на последний день третьей декады, заносится в см с точностью до целых;
- 17) норма по запасу воды в снежном покрове на последний день третьей декады, заносится в мм с точностью до целых.

В 14 строку заносится информация по 4 видам атмосферных явлений (их идентификационные номера) и коэффициенты степенного уравнения регрессии (a и b): туман – **10**, метель – **11**, гололед – **12**, изморозь - **13**. Как видно из таблицы 2, для станции 5392091 рассчитаны коэффициенты регрессии только для тумана и метели.

В РР01 для номера атмосферного явления отведено 2 байта, для коэффициентов a и b – по 3 байта. Порядок расчета коэффициентов степенного уравнения регрессии для контроля атмосферных явлений приведен в пункте 2.3 данного методического письма.

При занесении дробных чисел точка не ставится. В случае отсутствия каких-либо данных, место, отведенное для них, остается незаполненным. При отсутствии коэффициентов уравнений регрессии для атмосферных явлений в 14-ую строку заносится только координатный номер контролируемой станции.

Для каждой контролируемой станции должно быть отведено 14 строк независимо от наличия информации в них. Паспорт РР01 заполняется обязательно для всех станций независимо от метода контроля и наличия норм.

2.3 Коэффициенты уравнений регрессии для контроля метеорологических величин

В районах с редкой сетью, где не удастся подобрать необходимое количество влияющих станций, контроль проводится с использованием пространственно-временных связей (уравнения линейной регрессии) значений ряда метеорологических величин на отдельных станциях. Для конкретной метеорологической величины на контролируемой станции находится **станция-аналог** по этой же метеорологической величине (потому станция-аналог может быть одна, если она подходит для всех контролируемых метеорологических величин, либо несколько станций-аналогов для соответствующих метеорологических величин.).

Не всегда удастся подобрать станцию-аналог, особенно для изолированных станций в районах, где практически отсутствует метеорологическая сеть, в таких случаях контроль качества данных производится на основе временных внутростанционных связей одних метеорологических величин с другими. Основу таких связей составляют уравнения линейной регрессии между контролируемой величиной и **величиной-аналогом** на контролируемой станции.

При использовании временных внутростанционных связей между метеорологическими элементами, для которых предусмотрен расчет уравнения линейной регрессии, рекомендуется пользоваться таблицей 3. При этом следует иметь в виду, что некоторые связи устойчивы только для конкретных месяцев года.

Если контроль осуществляется с использованием уравнения линейной регрессии хотя бы по одному метеорологическому элементу (кроме атмосферных явлений), необходимо заполнить паспорт РР02 (массив «коэффициенты уравнений»), где содержится информация о станциях-аналогах или величинах-аналогах.

Важно отметить, что, если при контроле какой-либо величины используется ее величина-аналог, то для контроля этой величины-аналога нельзя использовать уже проконтролированную характеристику, а необходимо найти третью величину-аналог (к примеру, если для контроля температуры подстилающей поверхности использовались данные по температуре воздуха, то для контроля температуры воздуха уже нельзя использовать данные температуры подстилающей поверхности).

Таблица 3

Метеорологические величины-аналоги

Контролируемая величина \ Величина-аналог	Температура воздуха	Температура подстил. пов.	Относительная влажность	Дефицит насыщения	Парциальное давление	Скорость ветра	Кол-во общей облачности	Кол-во нижней облачности
Температура воздуха		+	+		+			
Температура подстил. пов-ти	+							
Относительная влажность	+				+	+		
Дефицит насыщения	+		+		+			
Парциальное давление	+		+	+				
Скорость ветра			+					+
Кол-во общей облачности								+
Количество нижней облачности			+			+	+	

Коэффициенты уравнений линейной регрессии (связи) необходимо обновить по тем же причинам, что и «нормы». Уравнение линейной регрессии в общем виде записывается как:

$$U_k = aF_a + b \quad (1),$$

где:

- U_k – рассчитанное значение контролируемой метеорологической величины на станции;
- F_a – значение метеорологической величины на станции-аналоге или значение метеорологической величины-аналога;
- a, b – коэффициенты уравнения линейной регрессии.

Коэффициенты a и b для каждого месяца и для каждой контролируемой величины рассчитываются за последние 15 лет (с 2001 по 2015 гг.) методом наименьших квадратов:

$$a = \frac{\sum_{i=1}^n (F_{ai} - \bar{F}_a)(F_{ki} - \bar{F}_k)}{\sum_{i=1}^n (F_{ai} - \bar{F}_a)^2} \quad (2),$$

$$b = \bar{F}_k - a\bar{F}_a \quad (3),$$

где:

- F_{ki} – средние месячные значения контролируемой величины на контролируемой станции;
- F_{ai} – средние месячные значения контролируемой величины на станции-аналоге или средние месячные значения величины-аналога на контролируемой станции;
- i – годы с 2001 по 2015 гг.;
- n – 15;
- \bar{F}_k – среднее многолетнее месячное значение контролируемой величины на контролируемой станции за период 2001-2015 гг.;
- \bar{F}_a – среднее многолетнее месячное значение контролируемой величины на станции-аналоге или среднее многолетнее месячное значение величины-аналога на контролируемой станции за период 2001-2015 гг.

Перед расчетом коэффициентов a и b из выборки **исключаются** аномальные значения членов F_{ki} и F_{ai} , разность которых превышает утроенное значение среднего квадратического отклонения (σ) средней разности по выборке (пример расчета представлен в таблице 4б).

Кроме того, следует рассчитать коэффициент корреляции между метеорологической величиной на контролируемой станции и соответствующей метеорологической величиной на станции-аналоге или величиной-аналогом. Уравнением связи для контроля характеристики можно пользоваться, если коэффициент корреляции не менее 0,7.

При контроле атмосферных явлений, таких как *туман, метель, изморозь, гололед*, используется уравнение связи между числом дней с контролируемым явлением на станции и его суммарной продолжительностью:

$$y_i = a \cdot x_i^b \quad (4),$$

где y_i – число дней с явлением, x_i – продолжительность явления в часах, a и b – заранее рассчитанные коэффициенты уравнений связи.

$$a = \exp \frac{\sum_{i=1}^n \ln y_{i\Phi} \sum_{i=1}^n \ln x_i - \sum_{i=1}^n \ln x_i \sum_{i=1}^n \ln y_{i\Phi} \ln x_i}{n \sum_{i=1}^n (\ln x_i)^2 - (\sum_{i=1}^n \ln x_i)^2} \quad (5),$$

$$b = \frac{n \sum_{i=1}^n \ln y_{i\Phi} \ln x_i - \sum_{i=1}^n \ln x_i \sum_{i=1}^n \ln y_{i\Phi}}{n \sum_{i=1}^n (\ln x_i)^2 - (\sum_{i=1}^n \ln x_i)^2} \quad (6),$$

где $y_{i\Phi}$ – фактическое число дней с явлением, n – общее число рассматриваемых случаев. Уравнение регрессии для атмосферных явлений (метель, туман, изморозь, гололед) рассчитывается одно для всех месяцев года, и его коэффициенты заносятся с точностью до сотых в 14 строку паспорта PP01.

Для реализации расчета коэффициентов *степенного* уравнения связи (уравнения 5, б) рекомендуется использовать разработанные в ФГБУ «ГГО» программы расчета REGR1, REGR2 (прилагаются к письму необходимые файлы.exe и их полное описание в папке REGR) или с помощью стандартного приложения Excel (для этого необходимо построить график, где по оси абсцисс (x) задаются значения продолжительности явления в часах, по оси (y) число дней с явлением; далее вывести на график степенное уравнение тренда, где коэффициенту a соответствует параметр перед множителем x , коэффициенту b соответствует степень, в которую возводится множитель x).

2.3.1 Программы REGR1, REGR2 для расчета коэффициентов уравнения регрессии

Для автоматизированного расчета уравнений регрессии необходимо сохранить папку REGR (которая прикреплена к письму в качестве электронного приложения) в директории `c:\METEOWIN\REGR\`. В папке REGR находятся файлы REGR1.exe и REGR2.exe, а также REGR1.doc и REGR2.doc.

Программы REGR1, REGR2 предназначены для расчета коэффициентов уравнений линейной регрессии (уравнения 2, 3) и степенной регрессии (уравнения 5, 6). REGR1 реализована для стандартной подготовки исходной информации (т.е. исходная информация заносится строго по подготовленной форме, описанной в файле REGR1.doc). REGR2 реализована для произвольной подготовки исходной информации (полное описание дается в файле REGR2.doc). Обе программы рассчитывают необходимые для анализа статистические характеристики исходных данных наблюдений, автоматически исключают из расчета аномальные значения, рассчитывают коэффициенты уравнений регрессии как для метеорологических величин, так и для атмосферных явлений, коэффициенты корреляции, среднеквадратические отклонения и т.д.

В стандартный набор программ «ПЕРСОНА-МИС» данные программы не входят, поэтому при каждом обновлении системы «ПЕРСОНА-МИС» необходимо копировать папку REGR в обновленную версию.

2.3.2 Расчет коэффициентов линейного уравнения регрессии в EXCEL

Коэффициенты линейного уравнения регрессии также возможно рассчитать в стандартном приложении Excel.

На начальном этапе необходимо подготовить массив исходных данных для метеорологических величин, контроль которых проводится методом уравнений регрессии. Исходные данные представляют собой средние месячные значения метеорологической величины на контролируемой станции и на станции-аналоге; при внутристанционном контроле – среднее месячное значение контролируемой величины и соответствующее значение величины-аналога на этой же станции. Желательно исходные данные (средние месячные значения метеовеличин) рассчитывать с точностью до сотых.

Если метеорологическая характеристика контролируется по уравнениям регрессии (кроме атмосферных явлений), то для каждого месяца должны быть подготовлены массивы исходных данных и для каждого месяца получены уравнения регрессии (12 шт.).

В качестве примера в таблице 4а представлен массив исходных данных для внутристанционного контроля температуры подстилающей поверхности по температуре воздуха для июля.

Таблица 4а

Средние месячные значения температуры подстилающей поверхности
и температуры воздуха, июль

Год	t подст.пов-ти (контролируемая величина, U_k)	t воздуха (величина-аналог, F_a)
2001	16,50	14,18
2002	22,94	19,16
2003	20,10	16,70
2004	17,98	16,40
2005	24,46	20,46
2006	18,16	15,20
2007	21,60	16,78
2008	15,86	12,98
2009	17,36	14,73
2010	19,16	15,37
2011	19,75	17,02
2012	19,57	15,97
2013	18,59	15,06
2014	17,79	14,98
2015	17,73	15,48

Коэффициент корреляции составляет 0,95, что позволяет судить о прочности связи между температурой воздуха и температурой подстилающей поверхности в июле.

Для того чтобы воспользоваться инструментом регрессионного анализа, встроенного в Excel, необходимо активировать надстройку **Пакет анализа**. Найти ее можно, перейдя по вкладке **Файл** → **Параметры**, в появившемся диалоговом окне **Параметры Excel** переходим во вкладку **Надстройки**. В поле **Управление** выбираем **Надстройки Excel** и щелкаем **Перейти**. В появившемся окне ставим галочку напротив **Пакет анализа**, нажимаем «**ОК**». Во вкладке **Данные** в группе **Анализ** появится новая кнопка «**Анализ данных**» (рис.2).

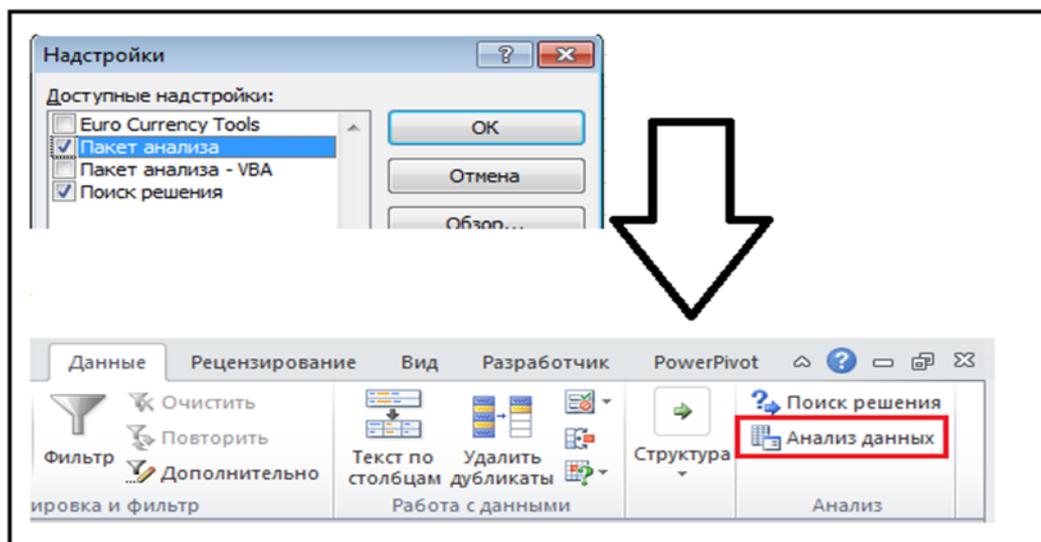


Рисунок 2– Активация кнопки Анализ данных.

Далее в появившейся кнопке «**Анализ данных**» необходимо выбрать **Регрессия** и нажать «**ОК**» (рис. 3)

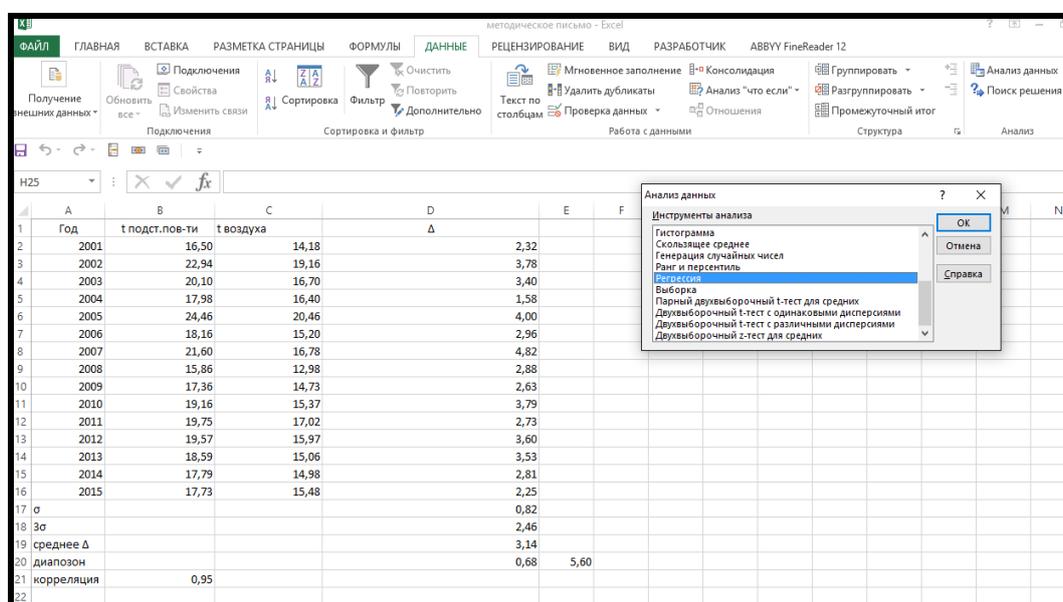


Рисунок 3 – Вызов подпрограммы «Регрессия».

В появившемся диалоговом окне (рис.4) необходимо указать входные параметры для расчета. Входной интервал **Y** – средние месячные значения контролируемой метеорологической характеристики за 15 лет (соответствует параметру U_k в уравнении 1), например, температура подстилающей поверхности за июль. Входной интервал **X** (соответствует параметру F_a в уравнении 1) – средние месячные значения

метеорологической характеристики за 15 лет, наблюдаемые на станции-аналоге (или при внутростанционном контроле - величина-аналог, например, температура воздуха за июль). Выбираем уровень надежности 95%, выбираем удобный вариант в графе «Параметры вывода» и необходимые характеристики в графе «Остатки» (дополнительно можно выводить график остатков, график подбора, график нормальной вероятности и т.д., если это требуется для дальнейшего анализа).

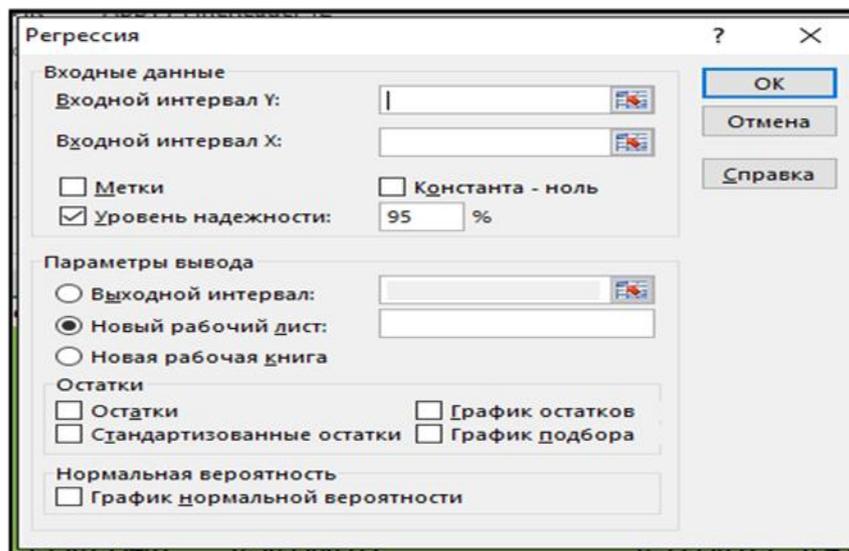


Рисунок 4 – Диалоговое окно «Регрессия».

После работы подпрограммы формируется новый лист с результатами расчета (рис.5). В качестве примера представлены результаты расчета регрессионного анализа между средними месячными значениями температуры воздуха (величина-аналог) и температуры подстилающей поверхности (контролируемая величина) за 15 лет (рис.5). Красным выделены коэффициенты уравнения регрессии, рассчитанные для данного конкретного месяца и конкретной станции. Рассчитанные коэффициенты уравнения регрессии округляются до сотых.

Таким образом, полученное уравнение регрессии для контроля температуры подстилающей поверхности в июле примет вид:

$$U_k = 1,19F_a + 0,09$$

Соответственно для каждого месяца аналогичным способом могут быть получены уравнения регрессии для каждой необходимой метеорологической характеристики.

Для исключения из выборки сомнительных данных необходимо рассчитать разность (Δ) между фактическими значениями контролируемой величины и соответствующими расчетными значениями контролируемой величины по уравнению регрессии (значения U_p и Δ выводятся дополнительно в «Выводах итогов» (рис.6)). Далее рассчитывается средняя разность ($\bar{\Delta}$) и среднее квадратическое отклонение (σ_{Δ}). Затем находятся допустимые пределы Δ : $[\bar{\Delta} - 3\sigma; \bar{\Delta} + 3\sigma]$ и из выборки исключаются те значения, Δ которых не входят в полученный интервал и коэффициенты уравнения регрессии пересчитываются заново без учета сомнительных значений.

1	ВЫВОД ИТОГОВ		
2			
3	<i>Регрессионная статистика</i>		
4	Множественный R	0,949901848	коэффициент корреляции
5	R-квадрат	0,90251352	
6	Нормированный R-квадрат	0,894799176	
7	Стандартная ошибка	0,763495437	
8	Наблюдения	15	
9			
10	<i>Дисперсионный анализ</i>		
11		<i>df</i>	<i>SS</i>
12	Регрессия	1	69,99697133
13	Остаток	13	7,578028673
14	Итого	14	77,575
15			
16		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>
17	Y-пересечение	0,094599338	1,751893053
18	Переменная X 1	1,189882355	0,108585253
19			
20			
21			
22			
23			

Рисунок 5 – Вывод результатов выполнения регрессионного анализа.

<i>Дисперсионный анализ</i>				
		<i>df</i>	<i>SS</i>	<i>MS</i>
Регрессия		1	69,99697133	
Остаток		13	7,578028673	
Итого		14	77,575	
		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>
Y-пересечение		0,094599338	1,751893053	
Переменная X 1		1,189882355	0,108585253	
ВЫВОД ОСТАТКА				
	<i>Наблюдение</i>	<i>Предсказанное Y</i>	<i>Остатки</i>	<i>Стандартные</i>
	1	16,96713113	-0,467131133	-
	2	22,89274526	0,047254738	-
	3	19,96563467	0,134365332	-
	4	19,60866996	-1,628669962	-
	5	24,43959232	0,020407677	-
	6	18,18081114	-0,020811135	-
	7	20,06082526	1,539174743	-
	8	15,53927231	0,320727693	-
	9	17,62156643	-0,261566429	-
	10	18,38309114	0,776908864	-
	11	20,34639702	-0,596397022	-
	12	19,09702055	0,472979451	-
	13	18,01422761	0,575772394	-
	14	17,91903702	-0,129037017	-
	15	18,51397819	-0,783978195	-

Рисунок 6 – Вывод значений контролируемой величины по уравнению регрессии (U_p) и разности между фактическими значениями контролируемой величины и значениями, рассчитанными по уравнению регрессии (Δ).

В приведенном примере (таблица 4б) все Δ входят в полученный интервал, соответственно все значения величин используются для расчета коэффициентов a и b и пересчет коэффициентов уравнения регрессии не требуется.

Таблица 4б

Средние месячные значения температуры подстилающей поверхности
и температуры воздуха, июль

Год	t подст.пов-ти (контролируемая величина, U_k)	t воздуха (величина-аналог, F_a)	t подст.пов-ти* (контролируемая величина, рассчитанная по уравнению связи, U_p)	$\Delta =$ $(U_k - U_p)$
2001	16,50	14,18	16,97	-0,47
2002	22,94	19,16	22,89	0,05
2003	20,10	16,70	19,97	0,13
2004	17,98	16,40	19,61	-1,63
2005	24,46	20,46	24,44	0,02
2006	18,16	15,20	18,18	-0,02
2007	21,60	16,78	20,06	1,54
2008	15,86	12,98	15,54	0,32
2009	17,36	14,73	17,62	-0,26
2010	19,16	15,37	18,38	0,78
2011	19,75	17,02	20,35	-0,60
2012	19,57	15,97	19,10	0,47
2013	18,59	15,06	18,01	0,58
2014	17,79	14,98	17,92	-0,13
2015	17,73	15,48	18,51	-0,78
Средняя разность $\bar{\Delta}$				0,00
σ_{Δ}				$\pm 0,74$
$3 \cdot \sigma_{\Delta}$				$\pm 2,22$
Диапазон $[\bar{\Delta} - 3\sigma; \bar{\Delta} + 3\sigma]$			$[-2,22; 2,22]$	
Корреляция			0,95	

В качестве примера заполнения массива «коэффициенты уравнений», приведена станция Тыгда с координатным номером 5312641 (таблица 5). В таблице 5 рассмотрен случай заполнения паспорта РР02, когда контроль осуществляется по уравнениям связи со станциями-аналогами.

Паспорт «коэффициенты уравнений» для станции Тыгда (РР02)

№строки														
1	11	5312641	25	1										
2	5312641	1	5352581	101	-9	5352581	101	53	5352581	82	1211	5352581	81	5
3			5352581	96	1	5282601	103	38	5352581	91	-19			
4	5312641	2	5352581	97	-102	5352581	91	-315				5352581	99	2
5			5352581	84	13	5282601	82	68	5352581	94	-46			
6	5312641	3												
7														
8	5312641	4	5352581	99	29	5352581	97	31	5352581	124-1180	5352581	102	5	
9			5352581	100	13				5352581	90	27	5352581	116	17
10	5312641	5	5352581	95	79	5352581	109	-17	5352581	58	2435	5352581	81	127
11			5352581	83	139				5352581	73	177	5352581	129	18
12	5312641	6	5352581	95	102	5352581	92	199	5352581	75	1679	5352581	87	105
13			5352581	86	193	5282601	68	75	5352581	134	-240	5352581	140	21
14	5312641	7	5352581	93	157	5352581	71	754	5352581	67	2485	5352581	85	125
15			5352581	81	323				5352581	81	137	5352581	106	73
16	5312641	8	5352581	94	105	5352581	83	389	5352581	104	-269	5352581	110	-45
17			5352581	100	19	5282601	59	81	5352581	106	-37	5352581	132	17
18	5312641	9	5352581	96	49	5352581	98	51	5352581	93	610	5352581	98	9
19			5352581	93	73	5282601	102	24	5352581	101	-28	5352581	103	78
20	5312641	10	5352581	96	12	5352581	110	63	5352581	67	2156	5352581	81	47
21			5352581	91	46	5282601	81	72	5352581	68	139	5352581	116	0
22	5312641	11	5352581	101	-7	5352581	94	-152				5352581	93	6
23			5352581	101	5	5352581	78	-24	5352581	63	128			
24	5312641	12	5352581	103	57	5352581	99	-119	5352581	85	990	5352581	92	1
24			5352581	97	3	5282601	93	37	5352581	72	79			

Информация для каждой станции занимает 25 строк.

В первую строку записываются: вид информации (идентификатор паспорта РР02 – «11»), координатный номер станции, номер УГМС, признак контроля по уравнениям регрессии («1» – уравнение регрессии между одноименными величинами, используя станцию-аналог; «2» – уравнение регрессии между двумя различными метеорологическими величинами внутри одной станции, используя величину-аналог).

Коэффициенты уравнений регрессии для каждого месяца занимают **две строки**.

В первую строку для каждого месяца заносится:

- координатный номер контролируемой станции;
- номер месяца;
- координатный номер станции-аналога, с которой найдена связь по температуре воздуха;
- коэффициент a для температуры воздуха, с точностью до сотых, с учетом знака (например, коэффициент $a = 1,01$ запишется как «101»);
- коэффициент b для температуры воздуха (например, коэффициент $b = -0,09$ запишется как « -9»).

Между координатным номером станции-аналога и коэффициентами a и b пробелы не ставятся. Пробелом разделяется информация для различных метеорологических величин.

Далее аналогичная информация для других метеорологических величин заносится в РР02 в следующем порядке:

- температура поверхности почвы,
- относительная влажность воздуха,
- дефицит насыщения водяного пара.

Во вторую строку для каждого месяца, начиная с 12-го байта, заносятся - координатные номера станций-аналогов и коэффициенты уравнений регрессии для следующих метеорологических величин:

- парциальное давления водяного пара,
- скорость ветра,
- количество общей облачности,
- количество нижней облачности.

PP02 «коэффициенты уравнений» заполняется только для станций, на которых хотя бы одна метеорологическая величина (кроме атмосферных явлений) контролируется по уравнениям регрессии.

При занесении дробных чисел точка не ставится, при отсутствии каких-либо данных, место, отведенное для них, остается незаполненным.

3 Особенности проведения пространственного контроля результатов измерений АМС.

Для обеспечения достоверности метеорологической информации, поступающей от автоматических метеорологических станций (АМС) без персонала, необходимо обеспечить регулярное проведение всех уровней контроля качества результатов измерений АМС.

В настоящее время ВНИИГМИ-МЦД разработано программное обеспечение, интегрированное в ПО ПЕРСОНА МИС, по формированию режимной (климатической) части информации АМС на основе данных оперативных сообщений КН-01, что позволяет организовать автоматизированный контроль данных АМС.

Проведение автоматизированного пространственного контроля результатов измерений АМС в силу непродолжительных рядов наблюдений имеет свои особенности:

- контроль данных *по атмосферному давлению* должен проводиться для АМС постоянно независимо от длины ряда наблюдений, поскольку для проведения контроля атмосферного давления не требуется наличие норм;

- в случае, если АМС функционирует менее 10 лет, то для организации контроля ее данных необходимо подобрать ближайшую длиннорядную репрезентативную станцию-аналог, расположенную на равнинной территории;

- если АМС расположена в районе с густой метеорологической сетью и станция-аналог находится в однородных с АМС физико-климатических условиях, на расстоянии не более 50-60 км, то в качестве норм АМС рекомендуется использовать нормы станции-аналога и контролировать данные АМС методом интерполяции. Станцию-аналог не следует выбирать в качестве влияющей для этой АМС;

- если АМС расположена в районе с редкой метеорологической сетью, но станция-аналог находится в однородных с АМС физико-климатических условиях, то контроль данных АМС осуществляется методом уравнений регрессии, полученных для станции-аналога (в PP02 АМС указываются коэффициенты, полученные для внутростанционного контроля на станции-аналоге);

- при накоплении 10 лет непрерывных наблюдений на АМС следует рассчитать для нее *промежуточные* нормы и коэффициенты уравнений регрессии, внести изменения в PP01 и PP02, и проводить контроль традиционным образом.

Заключение

Качественное проведение пространственного контроля режимной метеорологической информации позволяет не только обеспечить пригодность ее для дальнейшего использования, но и дает потребителю гарантию качества всех данных.

Проведение работ по пересмотру влияющих станций, правильному выбору признака контроля, расчету и обновлению месячных «норм» станций и коэффициентов уравнений регрессии за рекомендованные периоды существенно повысит эффективность пространственного контроля, а, соответственно, и качество данных наблюдений.

Руководству УГМС следует учитывать, что с ростом количества АМС на территории, подведомственной УГМС, существенно увеличивается объем поступающей метеорологической информации, а соответственно возрастает нагрузка на специалистов отделов/групп метеорологии по обеспечению обработки и проведению контроля качества результатов наблюдений.

Отделам метеорологии УГМС (ЦГМС) необходимо включить работы по обновлению вспомогательных материалов для пространственного контроля в планы отделов и выполнить их в течение 2017 года. Методическое сопровождение и контроль за выполнением указанных работ будут проводиться ФГБУ «ГГО», как головной организацией в системе Росгидромета по научно-методическому руководству государственной наблюдательной метеорологической сетью.

Зав. Методическим отделом ФГБУ «ГГО»



С.Ю. Гаврилова

м.н.с. Методического отдела ФГБУ «ГГО»



Л.И. Коломеец