

И.Л. Галкина<sup>1</sup>, В.В. Зуев<sup>2</sup>, Г.М. Крученицкий<sup>1</sup>

## Использование электронных таблиц Excel для изучения озоносферных процессов: архивация, анализ, моделирование

<sup>1</sup> Центральная аэрологическая обсерватория, г. Долгопрудный Московской обл.

<sup>2</sup> Институт оптики атмосферы СО РАН, г. Томск

Поступила в редакцию 4.04.2003 г.

Обсуждаются возможности электронных таблиц Excel применительно к задачам архивации, анализа и доведения до пользователя различных экспериментальных данных, используемых при изучении различных аспектов озоновой проблемы. Показано, что применение Excel позволяет в рамках единого подхода обеспечить архивацию, сравнение и визуализацию результатов разнообразных наземных, балонных и спутниковых измерений, контроль качества работы озонометрической сети, решение разнообразных задач, связанных с обнаружением и диагностикой озоновых аномалий. Затронуты вопросы взаимодействия с другими пакетами прикладных программ и средствами управления базами данных, а также построения регрессионных моделей.

Изучение озоносферных процессов, как и атмосферных вообще, обладает рядом особенностей по сравнению с традиционным подходом, применяемым в естествознании при исследовании того или иного класса явлений. Это, прежде всего, обусловлено следующими специфическими, но традиционными для геофизики причинами:

- наличием большого числа иных процессов, оказывающих влияние на изучаемый, и обусловленных этим обстоятельством трудностей при идентификации повторяемости;
- невысокой (по сравнению с большинством других разделов физики) точностью измерений параметров изучаемого процесса;
- наличием большого числа обратных связей, охватывающих изучаемый и иные процессы, что сильно затрудняет как корректное определение перечня подлежащих измерению параметров в ходе натурных экспериментов, так и причинно-следственную диагностику происходящего.

Отмеченные особенности приводят к тому, что в атмосферных исследованиях существенно возрастает (по сравнению с другими разделами физики) потребность в наличии разнообразных баз данных, роль статистического моделирования и объем экспериментального материала, необходимого для тестирования тех или иных закономерностей, которые установлены в ходе изучения геофизического процесса.

Появление и широкое распространение в последние десятилетия персональных компьютеров и разнообразного программного обеспечения к ним позволяют существенно повысить возможности корректного проведения геофизических исследований. К сожалению, зачастую эти возможности использу-

ются для построения так называемых «численных моделей», содержащих сотни, тысячи и даже десятки тысяч плохо определенных констант (химических, оптических, газодинамических и др.), не допускающих качественного анализа и требующих огромных затрат машинного времени. По нашему мнению, альтернативой такому подходу является использование средств вычислительной техники по прямому назначению, т.е. для решения следующих задач:

- создания обширных и удобных в использовании баз данных;
- повышения соотношения сигнал-шум в рядах геофизических наблюдений;
- автоматизации рутинных вычислительных и информационно-поисковых операций;
- привлечения обширных (по возможности исчерпывающих) данных измерений к тестированию аналитических результатов или результатов, полученных на небольшом экспериментальном материале.

В настоящей статье мы стремились продемонстрировать возможности одного из самых массовых программных продуктов — электронных таблиц Excel [1] — для реализации такого подхода применительно к задачам изучения поведения озона в атмосфере. Сразу оговоримся: статья не претендует на исчерпывающее описание возможностей Excel применительно к озоновой проблеме. Такие возможности если не безграничны, то настолько велики, что для их детального описания потребовалась бы обширная монография. Наша цель гораздо скромнее — привлечь внимание коллег к использованию мощного и эффективного программного средства в повседневной практической работе.



Файл формата TOMS Grid легко усваивается рабочей книгой Excel с помощью одной строки VBA кода:

```
Workbooks.OpenText FileName:=path, Origin := xlWindows, StartRow:=4, DataType:=xlFixedWidth, _
FieldInfo:=Array(Array(0, 1), Array(4, 1), Array(7, 1),
Array(10, 1), Array(13, 1), Array(16, 1), _
Array(19, 1), Array(22, 1), Array(25, 1), Array(28, 1),
Array(31, 1), Array(34, 1), Array(37, 1), Array(40, 1), _
Array(43, 1), Array(46, 1), Array(49, 1), Array(52, 1), _
Array(55, 1), Array(58, 1), Array(61, 1), Array(64, 1), _
Array(67, 1), Array(70, 1), Array(73, 1), Array(76, 1))
```

Скопировав рабочий лист с данными в формате TOMS Grid на первый рабочий лист книги Excel, между первым и вторым рабочими листами которой установлена (заранее и программно) связь, разворачивающая данные в матрицу (см. Приложение 1), приходим к формату данных, которые легко интерпретировать, визуализировать и сравнивать с данными из других источников, например с данными наземных станций. Вид второго листа рабочей книги, на котором широта развернута вдоль столбцов, а долгота вдоль строк, приведен на рис. 2.

Такая развертка (транспонированная по отношению к обычной меркаторской проекции) применена потому, что число элементов разрешения по

долготе (288) превышает максимальное число столбцов на рабочем листе Excel (256). Приведенный пример иллюстрирует эффективность применения для целей быстрого просмотра данных, содержащихся в БД, такого мощного средства Excel, как установление связей между ячейками различных рабочих листов (в том числе и листов, находящихся в различных рабочих книгах).

На рис. 2 видны помеченные ячейки, в которых расположены станции наземной сети. Эти ячейки снабжены комментариями, содержащими имена станций. Визуализация глобального поля ОСО встроенными средствами Excel проиллюстрирована на рис. 3.

Графические возможности Excel позволяют строить соответствующие карты, благодаря десяткам цветов, что делает возможности разрешения по динамическому диапазону значений ОСО практически неограниченными. Меняя в приведенной выше строке кода значение строки path, указывающей путь к текстовому файлу в формате TOMS Grid, можно быстро и эффективно осуществлять как визуальный просмотр эволюции глобального распределения общего содержания озона, так и необходимые пользователю выборки его значений. Так, на рис. 4 проиллюстрированы результаты такой выборки, сделанной для оценки качества работы наземных станций Одесса и Киев во втором квартале 2002 г.

Рис. 2. Данные TOMS в виде транспонированной меркаторской матрицы

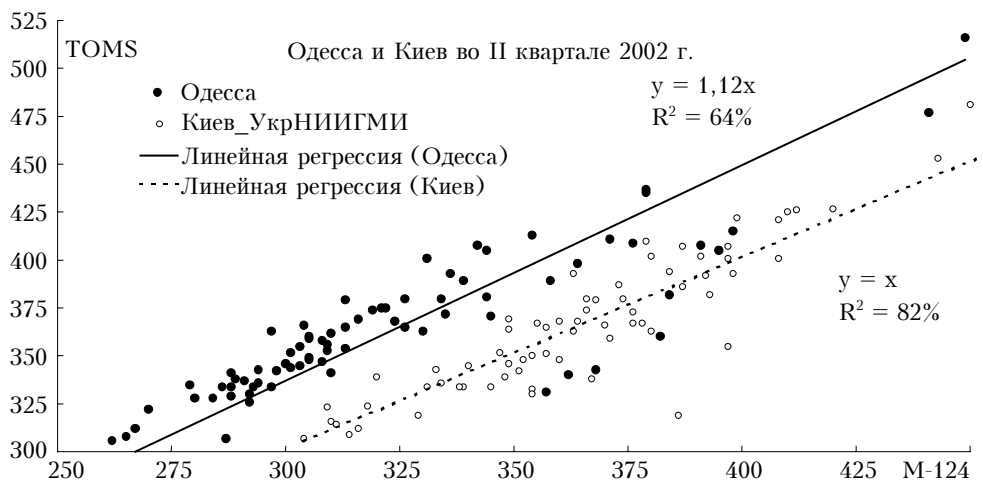


Рис. 4. Сравнение качества работы наземных станций Киев и Одесса

Рис. 4 и приведенные на нем расчетные параметры, характеризующие качество работы станций, выполнены встроенными средствами Excel в режиме выполнения макросов, т.е. замена списка станций (Одесса, Киев) на любой другой выполняется автоматически, и результаты сравнения получаются практически немедленно. Из рисунка видно, что качество работы ст. Киев существенно выше, так как приведенный коэффициент регрессии наземных и спутниковых данных практически идеален (равен 1), и разброс данных, определяемый коэффициентом детерминации регрессии  $R^2$  от единицы, гораздо меньше. (Вообще говоря, нельзя для любого сезона и любой станции требовать тождественности наземных и спутниковых данных не только из-за наличия погрешностей в спутниковых измерениях, но и ввиду того обстоятельства, что измеряются разные величины: со спутника полуденное (по местному времени) значение ОСО, а на наземной станции — среднеедневное. Однако отличие коэффициента регрессии от 1 на 12% свидетельствует о явном занижении результатов при наземных измерениях).

Применительно к возможностям, которые предоставляет Excel при архивации и визуализации спутниковых данных об ОСО, легко осуществить необходимое пространственное и временное осреднение первичных данных, необходимость в котором возникает ввиду следующих обстоятельств:

- малая высота орбиты КА «Earth Probe», на котором в настоящее время функционирует аппаратура TOMS, данные об ОСО в экваториальной зоне и отчасти в умеренных широтах содержат пропуски (см. рис. 3);
- отсчеты в соседних ячейках TOMS Grid в умеренных и приполярных широтах дублируют друг друга;
- для анализа долговременных изменений в состоянии глобального распределения ОСО пространственное разрешение, соответствующее размерам ячейки TOMS Grid, является заведомо избыточным.

Возможности Excel по осреднению данных TOMS демонстрирует рис. 5, для получения которого ежемесячно осредненные данные об ОСО пространственно осреднены с шагом  $3^\circ$  по широте и  $5^\circ$  по долготе. Осреднение осуществлено с помощью программно реализованного установления связей между листами рабочей книги с помощью встроенной функции Excel, образец которой приведен ниже.

```
ЕСЛИ(СЧИТАТЬПУСТОТЫ(TrMatr!R[183]
C[102]:R[186]C[104])>7,"",
CPЗНАЧ(TrMatr!R[183]C[102]:R[186]C[104]))
```

Приведенная функция осуществляет осреднение значений матрицы из  $3 \times 4$  ячеек TOMS Grid, если в этой матрице имеется не менее 8 отсчетов. Программно связь между листами устанавливается аналогично тому, как это сделано в макросе, исходный код которого приведен в Приложении 1. Создание топографической основы на листе диаграммы, приведенной на рис. 5, осуществлено с помощью макроса, исходные коды которого даны в Приложении 2. В случае необходимости (например, по издательским требованиям) построить черно-белую карту с высоким разрешением по динамическому диапазону можно, например, осуществить построение ее в MathCad, импортировав в него данные Excel. Пример такой карты приведен на рис. 6.

Перейдем теперь к архивации и анализу с помощью Excel данных о вертикальном распределении озона (ВРО). Эти данные также делятся на результаты балонных и спутниковых измерений. Существуют три формата, используемые для архивации данных о ВРО — это форматы WOUDC (новый и старый) и формат NASA (аэрокосмического агентства США) и данные в обоих этих форматах широко представлены на диске [3]. Все три формата легко усваиваются Excel, который представляет существенные преимущества при поиске конкретных профилей ВРО или при необходимости получения любой дополнительной информации о них.

Глобальное распределение ОСО в октябре 1998 г.

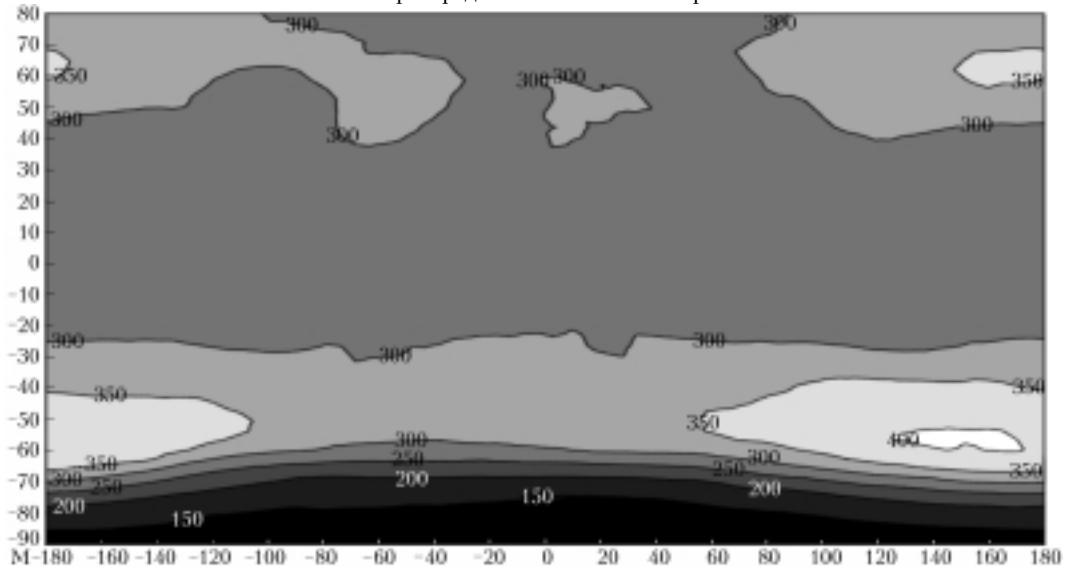


Рис. 6. Пример визуализации данных аппаратуры TOMS, импортированных в MathCad из Excel

Поясним сказанное примерами. Так как системы радиозондирования, применяемые на мировой сети балонных измерений ВРО, не используют радиолокатор в первичных данных двух из трех упомянутых выше форматов, высотная привязка отсчетов отсутствует. Встроенные функции Excel позволяют не только без труда восстановить эту привязку с помощью вытекающего из барометрической формулы рекуррентного соотношения

$$h_{i+1} = h_i + \frac{(T_{i+1} + T_i)R}{2\mu g} \ln \frac{p_i}{p_{i+1}},$$

где  $h_i$  — высота  $i$ -го отсчета;  $T_i = 273,15 + t_i^0$ ;  $t_i^0$  —  $i$ -й отсчет температуры;  $p_i$  —  $i$ -й отсчет давления;  $R$  — универсальная газовая постоянная;  $\mu$  — молекулярный вес воздуха;  $g$  — ускорение земного притяжения, или рассчитать профили отношения смеси озона, его концентрации, накопленной концентрации (Integrated Ozone), потенциальной температуры и т.д., но и произвести любые модельные расчеты для всего архива данных, задав необходимую формулу в единственной ячейке соответствующего столбца и распространив ее на всю книгу с помощью метода Autofill. При этом результаты моделирования легко визуализируются стандартными средствами Excel. Например, в нашей работе [4] показано, что вероятность прямого разрушения молекулы озона при столкновении на высоте  $h$  с частицей полярных стратосферных облаков (ПСО)  $W$  связана с химическим потенциалом озона  $v$  соотношением

$$W = 1 - \operatorname{erf} \left( \sqrt{\frac{v}{RT}} \right) + 2 \sqrt{\frac{v}{\pi RT}} \exp \left( -\frac{v}{RT} \right),$$

где  $\operatorname{erf}$  — функция ошибок [5];  $T$  — абсолютная температура на высоте  $h$ ;

$$v = \mu_3 gh + R \int_0^h \frac{T}{p_3} dp_3;$$

$\mu_3$  — молекулярный вес озона;  $p_3$  — парциальное давление озона.

Из рис. 7 видно, что на рабочем листе Excel удобно рассчитать профили химического потенциала, условной вероятности  $W$  и интенсивность стока озона на частицы ПСО, которая пропорциональна произведению  $W$  на величину площади поверхности частиц ПСО в единице объема. Предполагая, что эта площадь, в свою очередь, пропорциональна отклонению абсолютной температуры от температурного порога существования ПСО  $\sim 200$  К, легко с точностью до множителя рассчитать высотные профили интенсивности стока озона. Один из таких профилей, совместно с профилем концентрации, приведен на рис. 8, выполненном штатными средствами визуализации Excel.

Из рис. 8 видно, что антикорреляция концентрации озона и интенсивности его стока на поверхность частиц ПСО ничуть не менее наглядна по сравнению с антикорреляцией концентрации озона и концентраций окиси хлора, представленной на известном рисунке из работы [6], который в свое время обошел всю мировую печать от журнала «Playboy» до газеты «Пионерская правда», запугивая простодушных читателей мифической фреоновой угрозой.

Здесь уместно подчеркнуть еще одно важное достоинство Excel. Все расчеты, выполненные в его рамках, легко проверяемы с точки зрения заложенных формул и исключают вычислительные ошибки. Т.е. результаты, основанные на моделировании средствами Excel, автоматически удовлетворяют общепринятому в естественных науках требованию повторяемости, которое практически полностью предано забвению в современных моделях атмосферных процессов [7].

Еще одним существенным удобством, предоставляемым пользователю электронными таблицами Excel, является наличие встроенной функции LINEST (в русском интерфейсе ЛИНЕЙН), которая позволяет вычислить коэффициенты линейной регрессии моделируемого ряда на набор регрессоров,

Рис. 7. Вид рабочего листа Excel с архивом данных баллонного озонзондирования

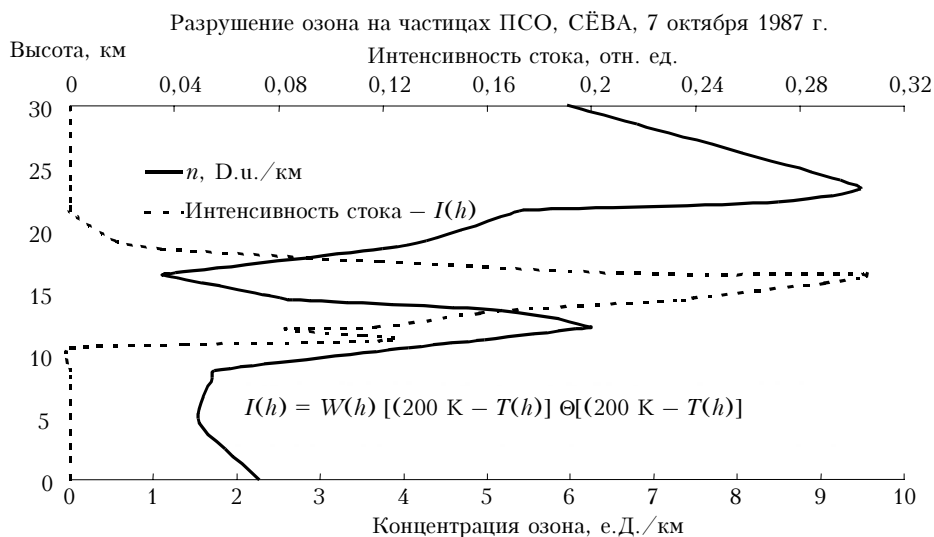


Рис. 8. Анतिकорреляция концентрации озона и интенсивности его стока на поверхность ПСО

включающий до 15 рядов. (Последнее ограничение не является существенным, один из способов обойти его изложен, например, в [8]). Кроме значений коэффициентов линейной регрессии функция LINEST возвращает и погрешности их вычисления (а также коэффициент детерминации модели, значение статистики Фишера и число степеней свободы). В сочетании с наличием в Excel функции, возвращаю-

щей значения *t*-статистики Стьюдента, это дает возможность определить уровень статистической значимости каждого из регрессоров и отобрать те из них, которые значимы на уровне, заданном пользователем. Разумеется, все эти операции легко автоматизируются макросами VBA. В качестве примера на рис. 9 приведен широтно-временной разрез сезонного хода ОСО по данным аппаратуры

TOMS, полученный регрессией зонально и помесячно осредненных рядов на статистически значимые (на уровне 95%) гармоники годового хода.

Этот рисунок также выполнен на основе импорта данных Excel в Mathcad для того, чтобы использовать более широкие возможности последнего пакета в части построения изолиний (как черных, так и цветных).

Excel позволяет также успешно производить поисковые операции по достаточно сложным критериям. Например, при анализе профилей ВРО и вертикального распределения температуры (ВРТ) важно выделить высоту тропопазы для того, чтобы отделить тропосферный озон, вызывающий в последнее время повышенный интерес исследователей, от остальных данных, а также для того, чтобы диагностировать тип воздушной массы (тропическая, полярная и т.д.).

Алгоритм выделения, основанный на определении [9], легко реализуется в виде макроса, спо-

собного обрабатывать данные обо всех выпусках зондов, содержащихся в архиве. (Тропауза определяется как первый уровень, лежащий не ниже изобарической поверхности 500 мбар, между которым и всеми более высокими уровнями в слое толщиной не менее 2 км средний вертикальный градиент температуры не превышает  $2^\circ/\text{км}$ .) Результаты автоматического определения высоты тропопазы легко тестировать методом экспертной оценки, выбирая с помощью встроенных функций Excel наибольшие и наименьшие значения высоты тропопазы. (Очень удобно для этой цели то, что Excel позволяет находить не просто наибольшую, но и вторую, третью и т.д. по величине высоту). На рис. 10 приведены ситуации, когда с помощью макроса Excel определены наибольшие и наименьшие значения высоты тропопазы на станции Hohenpeissenberg.

Из рис. 10 видно, что Excel успешно справляется с этой задачей даже в весьма нетривиальном

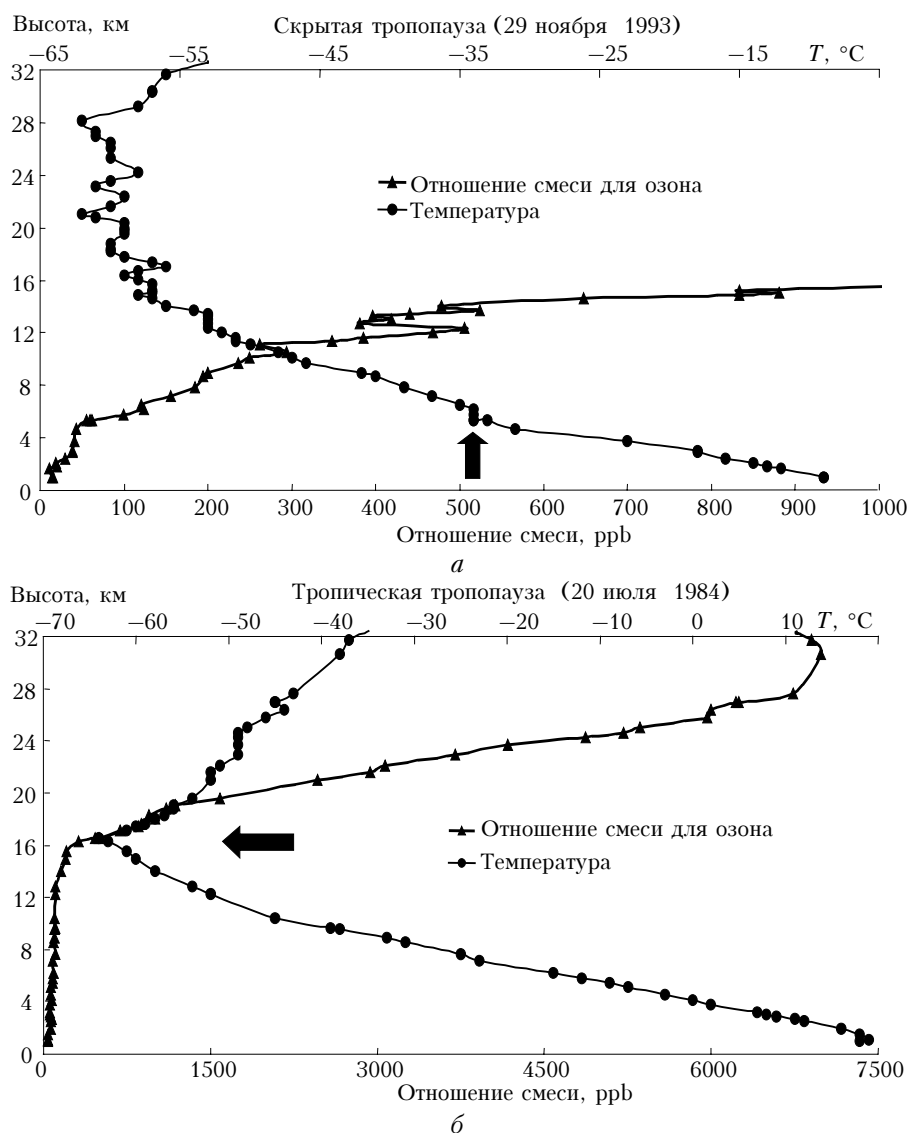


Рис. 10. Наименьшая (а) и наибольшая (б) высоты тропопазы для станции Hohenpeissenberg ( $48^\circ$  с.ш.,  $11^\circ$  в.д.) по результатам автоматической диагностики архива ее наблюдений макросом Excel

случае скрытой тропопаузы (рис. 10,а) и его диагностика подтверждается скачком градиента отношения смеси для озона.

Однако наиболее эффективно поисковые возможности Excel проявляются при работе с архивами спутниковых данных по вертикальному распределению метеорологических параметров и малых газовых составляющих атмосферы (МГСА). Используя встроенные фильтры Excel «Автофильтр» и «Пользовательский фильтр», а также метод «Сортировка», легко организовать данные в виде, удобном для анализа и визуализации. Фильтры позволяют оставить в рабочем листе только данные, удовлетворяющие заданным пользователем условиям (как по значениям параметров, так и по временной и географической привязке), а сортировка дает возможность упорядочить профильтрованный массив по убыванию или возрастанию тех или иных значений. В качестве примера

покажем широтно-высотный разрез отношения смеси озона по данным многоканального ИК-радиометра «CRISTA-2», производившего измерения с борта КА класса «Shuttle» в августе 1998 г. На рис. 11 приведен этот разрез, для получения которого кроме фильтров и сортировки использовалась встроенная функция Excel «Average», позволявшая легко получать зонально осредненные значения отношения смеси.

Разумеется, приведенные примеры далеко не исчерпывают возможности электронных таблиц Excel для задач архивации анализа и визуализации озонметрических данных. Однако мы надеемся, что они убедят наших коллег шире использовать в своей повседневной работе этот мощный и эффективный инструмент, который помимо повышения производительности труда специалистов существенно облегчит обмен данными и взаимную проверку получаемых научных результатов.

## ПРИЛОЖЕНИЕ 1

### Коды, устанавливающие связи между листами книги Excel для преобразования формата TOMS Grid к матрице отсчетов

Global Const TGFCOLUMNS = 25'    Число строк в широтном блоке формата TOMS Grid  
 Global Const TGLATES = 180'    Число широтных блоков  
 Global Const TGLONES = 288'    Число отсчетов (элементов разрешения) внутри широтного блока

Public Sub ToTrMatr()

'Преобразует формат TOMS Grid на первом рабочем листе в транспонированную матрицу на втором.

Dim lat As Integer, lon As Integer, ro As Integer, co As Integer, f As String

For lat = 1 To TGLATES

For lon = 1 To TGLONES

ro = 12 \* lat - 11 + (lon - 1) \ TGFCOLUMNS

co = (lon - 1) Mod TGFCOLUMNS + 1

f = Worksheets(1).name & "!R" & Trim(ro) & "C" & Trim(co)

f = "=" & "IF(" & f & "=0,""" & f & """)"

Worksheets(2).Cells(lon + 1, lat + 1).FormulaR1C1 = f

Next lon

Next lat

End Sub

## ПРИЛОЖЕНИЕ 2

Public Sub GeogrBG()

'Рисует на диаграмме ("Map") топографическую основу, долгота и широта точек,

'образующих которую, приведены в двух первых столбцах рабочего листа («GBG»)

Dim intRow As Integer, sngX As Single, sngY As Single, intI As Integer

Charts("Map").Activate

intRow = 1

sngSLat = Worksheets("GBG").Cells(1, 4).Value

sngNLat = Worksheets("GBG").Cells(1, 5).Value

sngKlon = Charts("Map").Axes(xlCategory).Width / 355

sngBlon = Charts("Map").Axes(xlCategory).Left + 177.5 \* sngKlon



```

sngKlat = Charts("Map").Axes(xlSeries).Height / (sngSLat - sngNLat)
sngBlat = Charts("Map").Axes(xlSeries).Top - sngNLat * sngKlat
Do While Worksheets("GBG").Cells(intRow, 1) <> ""
    intI = intI + 1
    If Abs(Worksheets("GBG").Cells(intRow, 1).Value) < 177.5 And _
        Worksheets("GBG").Cells(intRow, 2).Value > sngSLat And _
        Worksheets("GBG").Cells(intRow, 2).Value < sngNLat Then
        sngX = Worksheets("GBG").Cells(intRow, 1).Value * sngKlon + sngBlon - 0.5
        sngY = Worksheets("GBG").Cells(intRow, 2).Value * sngKlat + sngBlat - 0.5
        ActiveChart.Shapes.AddShape(msoShapeRectangle, sngX, sngY, 0.1, 0.1).Select
    If intI Mod 400 = 0 Then DoEvents
    End If
    intRow = intRow + 1
Loop
'sngX = sngBlat

ActiveChart.Shapes.AddLine(-177.5 * sngKlon + sngBlon, sngBlat, _
    177.5 * sngKlon + sngBlon, sngBlat).Select
Selection.ShapeRange.Line.Weight = 0.5
Selection.ShapeRange.Line.DashStyle = msoLineDash
ActiveChart.Shapes.AddLine(sngBlon, sngBlat + Worksheets("GBG").Cells(1, 4).Value *
sngKlat, _
    sngBlon, sngBlat + Worksheets("GBG").Cells(1, 5).Value * sngKlat).Select
Selection.ShapeRange.Line.Weight = 0.5
Selection.ShapeRange.Line.DashStyle = msoLineDash
Worksheets("GBG").Cells(1, 8).Value = 1
End Sub

```

1. Блатнер П., Ульрих Л., Кук К., Дик Т. Использование Microsoft Excel. М.; СПб.; Киев: Вильямс, 2000. 1024 с.
2. Гарнаев А. Самоучитель VBA. СПб.: БХВ, 1999. 503 с.
3. Ozone Data for the World. ODW CD#1. Environment. Canada, 1999.
4. Звягинцев А.М., Зуев В.В., Крученицкий Г.М., Ско-робогатый Т.В. О вкладе гетерофазных процессов в формирование весенней озоновой аномалии в Антарктиде // Исслед. Земли из космоса. 2002. № 3. С. 29–34.
5. Янке В., Эмде Ф. Таблицы функций. М.: Физматгиз, 1959. С. 120–122.
6. Anderson J.G., Brune W.H., Lloyd S.A. et al. // J. Geophys. Res. 1989. V. 94. P. 11480.
7. Галкина И.Л., Крученицкий Г.М. О принципиальных ограничениях в моделировании атмосферных процессов // Оптика атмосф. и океана (в печати).
8. Звягинцев А.М., Кадыгров Н.Е., Крученицкий Г.М. Регрессионный анализ зонально осредненных рядов спутниковых наблюдений за общим содержанием озона: регрессоры и тренды // Исслед. Земли из космоса. 2003. № 3 (в печати).
9. Наставление гидрометеорологическим станциям и постам. Вып. 4. Аэрологические наблюдения на станциях. Ч. IIIа. Л.: Гидрометеопиздат, 1973. 113 с.

***I.L. Galkina, V.V. Zuev, G.M. Kruchenitskii. Study of ozonospheric processes using MS Exel Tables: archiving, analysis, and simulation.***

The potentials of the MS Excel electronic tables in archiving, analysis, and presentation of experimental data employed in studying various ozone aspects are discussed. It is shown that the use of Excel enables, in the framework of a single approach, archiving, comparison, and visualization of the results of diverse ground-based, balloon- and satellite-borne measurements, ozonometer network quality control, and solution of various problems of ozone anomaly detection and diagnostics. Some issues of interaction with other applications and database management tools, as well as construction of regression models are considered.