

УДК 521.322:523.31-852, 551.510.535

## КЛАСТЕРНЫЙ АНАЛИЗ БАЗЫ ДАННЫХ ПАРАМЕТРОВ ОРБИТ ИСЗ

Л.С.Шакун, Н.И.Кошкин

Астрономическая обсерватория ОНУ им. И.И.Мечникова, Одесса, Украина  
*leonidserg08@yandex.ru, nikkoshkin@yahoo.com*

**ABSTRACT. Cluster analysis of database of orbit parameters of artificial satellites.** *L.Shakun, N.Koshkin.* The relational base of orbital parameters of near-Earth space objects (SO) is created. For 2007 it is led correlative and cluster analysis on variations of values  $B^*$  for 4.5 thousand of low-Earth orbit (LEO) objects. Clusters LEO with similar character of atmospheric drag are selected.

### Введение

Наличие собственного обширного каталога или доступ к иному открытому и обновляемому каталогу параметров орбит ИСЗ позволяет осуществлять глобальный мониторинг нейтральной составляющей атмосферы на высоте от 300 до 1000 км по торможению спутников. Украина не располагает собственным достаточным каталогом параметров орбит ИСЗ для решения такой задачи. В свободном доступе таких каталогов тоже нет, но ограниченный доступ к действительно актуальному каталогу параметров спутниковых орбит «Норад» (объединённая система аэрокосмической обороны США и Канады) позволяет продемонстрировать возможности и методику анализа подобных баз данных. Данный каталог ежедневно публикует параметры орбиты более 10000 космических объектов (КО). Распространение этой информации происходит через сеть Интернет в виде текстовых файлов (см., например, <http://www.celestrak.com/NORAD/elements/>), обновляемых иногда несколько раз в сутки. Каждый текстовый файл содержит элементы орбиты большинства каталожных КО, полученные накануне. Орбитальные параметры ИСЗ в каталоге принято распространять в виде списка двухстрочных наборов элементов TLE [1]. Такой формат представления информации удобен для применения, но неудобен для ретроспективного анализа. Основными недостатками такого способа хранения информации являются сложность извлечения исторических данных по интересующему объекту, а также контроль уникальности данных внутри каталога. Для решения

этих и ряда других задач нами была разработана реляционная база данных.

### Структура базы данных

Для сокращения времени для реализации реляционной базы данных параметров орбит ИСЗ была выбрана программа Microsoft Access. Эта программа обладает простым и удобным интерфейсом для разработки баз данных. Она широко известна и существует обширный круг людей умеющих работать с этой базой данных.

Каждый набор элементов орбиты TLE описывает промежуточную орбиту одного ИСЗ в соответствии с моделью SGP4/SDP4 [2]. В силу этого структура базы данных элементов орбиты ИСЗ (рис. 1) была выбрана в значительной степени в соответствии с форматом TLE.

Как видно из рисунка 1 для хранения элементов орбиты ИСЗ используется три таблицы.

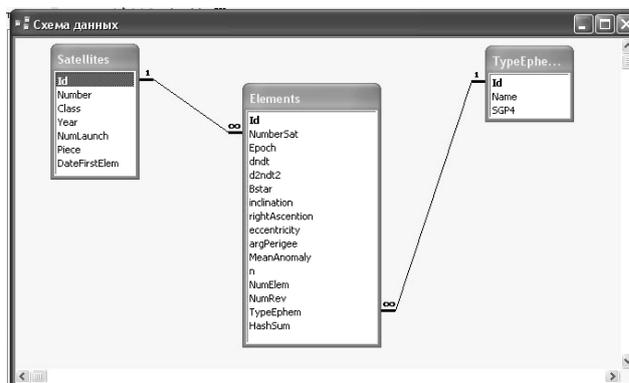


Рис. 1. Структура хранения данных параметров орбит ИСЗ в базе данных Access.

В таблице “Satellites” хранится перечень всех КО, когда либо сопровождавшихся и представленных в каталоге с 2005 года. Эта таблица содержит

внутренний уникальный номер спутника в нашей базе данных “Id”. Далее содержится уникальный номер объекта по каталогу «Норад» – “Number” и международный номер спутника в полях “Year”, “NumLaunch”, “Piece”. Представлена также эпоха первого набора элементов.

Таблица “Elements” хранит элементы орбит всех КО. Список полей таблицы эквивалентен полям двухстрочных элементов. Для указания формата используется поле “TypeEphem”. Поле таблицы “TypeEphem” может совпадать с обозначениями принятыми в каталоге «Норад», а может содержать и другие значения для обеспечения возможности хранения в этой базе данных элементов орбиты отличающихся от них. Например, эта база данных может хранить оскулирующие элементы орбиты КО или элементы других промежуточных орбит КО, чей перечень элементов совместим со списком элементов орбиты модели SGP4/SDP4.

Элементы орбиты в формате TLE часто обновляются, при этом они бывают действительно новыми, а иногда имеют минимальные различия в одной цифре или же совершенно не различаются. Эффективно проконтролировать уникальность всех полей таблицы “Elements” средствами программы Access невозможно. Поэтому для каждого набора элементов вычисляется контрольная сумма “HashSum”, которая с очень высокой вероятностью позволяет контролировать уникальность набора элементов орбиты ИСЗ.

#### Основные возможности реляционной базы данных элементов орбиты КО

Разработанная база данных элементов орбит КО позволяет извлекать, редактировать и добавлять данные в базу путём использования реляционных запросов на языке SQL. Запросы могут быть созданы средствами Microsoft Access или средствами других программ. Результаты запросов могут возвращаться в среду Microsoft Access, Microsoft Excel, Matlab, текстовый файл и т.д.

Нами разработана программа и система SQL запросов, позволяющих импортировать данные элементов орбит, список новых объектов из файлов, распространяемых в формате TLE. При этом в момент импорта происходит контроль уникальности элементов орбиты КО с целью недопущения дублирования данных. Разработана также программа, позволяющая осуществлять экспорт отобранных из базы данных элементов орбит КО в текстовые файлы, структурированные как двухстрочные элементы в формате TLE.

Для периода с марта 2005 года по июнь 2011 года наша база содержит около 30 млн. элементов орбит КО для более чем 13 тыс. объектов. Поиск и отбор объектов, их классификация по элементам орбиты в реальном времени отнимают нерационально много времени поэтому были разработаны запросы, рассчитывающие средние элементы КО на некоторый интервал, например, на год. Результаты этих запросов регулярно обновляются и хранятся в базе данных. Это

позволяет быстро извлекать наборы средних элементов и классифицировать КО по параметрам орбиты.

Распределение количества элементов орбит КО по годам можно увидеть в таблице 1. Основное количество архивных элементов орбиты приходится на 2005-2011 года. Видно, что количество объектов с 2005 по 2011 год неуклонно росло в основном за счёт роста числа осколков развалившихся ИСЗ.

Таблица 1. Распределение количества наборов элементов орбит в базе данных по годам.

Год	Кол. элементов орбит КО за год	Год	Кол. элементов орбит КО за год
2005	3012440	2009	5161416
2006	3901811	2010	5144625
2007	4418015	2011	3445324
2008	4309455		

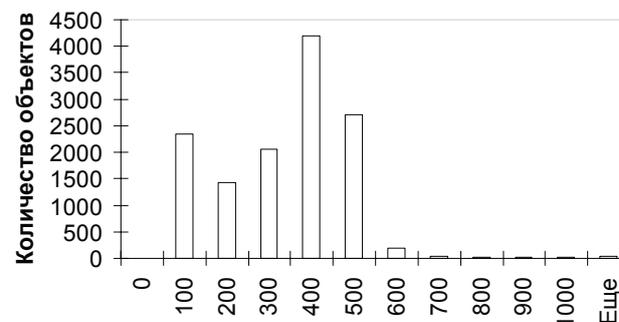


Рис. 2. Распределения количества объектов от среднего количества элементов орбиты за год

На рисунке 2 представлена гистограмма количества объектов в базе данных от среднего числа наборов параметров орбиты в течение года в период с 2005 по 2011 год. Видно, что характерное число новых наборов элементов орбиты для одного КО в год колеблется от 100 до 600, что соответствует обновлению элементов от одного раза в 3 дня до 2 раз в день.

#### Анализ движения ИСЗ в 2007 году

Нейтральная составляющая атмосферы на высотах до 1000 км над поверхностью земли оказывает тормозящий эффект на ИСЗ, движущиеся на этих высотах. В модели SGP4 параметр орбиты, отвечающий за учёт тормозящего влияния атмосферы, называется  $V^*$ . Ранее было нами показано [3], что для ряда ИСЗ, находящихся на орбитах близких к круговым, изменение параметра  $V^*$  хорошо коррелирует с индексом солнечной активности (F10.7). В тоже время было выявлено ряд КО на орбитах сходных с предыдущей группой, для которых изменение параметра  $V^*$  плохо коррелирует с индексами солнечной и геомагнитной активности. Причины такого поведения могут быть разными и

вызваны, например, особенностями формы КО и его вращения или характерными особенностями орбитальных параметров. Таким образом, возникает задача выявления групп КО, у которых зависимости изменения параметра торможения со временем сходны между собой.

Другая важная задача, требующая сортировки спутников по сходству их торможения, вызвана задачей оценки плотности нейтральной составляющей по известному параметру торможения ИСЗ  $V^*$ . Параметр  $V^*$  зависит не только от плотности атмосферы, но и от площади поперечного сечения, массы и баллистического коэффициента спутника. Существует очень незначительное количество спутников, для которых сечение постоянно, а баллистический коэффициент известен или может быть рассчитан теоретически. Масса спутника так же далеко не всегда известна. Для спутников с постоянной площадью сечения, известной массой и баллистическим коэффициентом, плотность атмосферы можно определять абсолютным методом. Для других спутников, чьи параметры неизвестны или могут изменяться со временем, необходимо контролировать их свойства через эталонные спутники, чьи характеристики известны хорошо. Таким образом, кластеризация спутников по сходству торможения может дать предварительные результаты для выбора семейств ИСЗ для относительной оценки плотности атмосферы.

Для выявления групп спутников со сходным характером торможения нами был выбран период с 1 января 2007 года по 31 декабря 2007 года. Из базы данных были отобраны все элементы орбиты для КО удовлетворяющих следующим условиям:

- В 2007 году КО имеет высоту перигея ниже 1000 км.
- КО не был запущен в 2007 году.
- Количество оценок параметров орбиты превышает количество дней в году, или, что тоже самое, среднее время между обновлениями элементов орбиты меньше суток.
- У КО параметр  $V^*$  изменяется в течение года.
- Объект не упал в первой половине 2007 года.

Этим условиям удовлетворило 4688 объектов. Очевидно, обработать такое количество объектов вручную не представляется возможным. Поэтому нами был разработан метод кластеризации объектов.

### Корреляционный анализ изменения торможения

Наиболее естественным параметром, характеризующим меру сходства двух отдельно взятых КО по их торможению, является коэффициент корреляции двух временных рядов параметра  $V^*$ , полученных в рамках модели SGP4.

Элементы орбит представлены в каталоге на момент прохождения КО восходящего узла орбиты и уточняются (обновляются) с некоторой не строгой периодичностью. Таким образом, для каждого КО значения параметра  $V^*$  приводятся на различные моменты времени и с различной периодичностью. Более того, регулярно встречаются случаи когда на

значительном промежутке времени (порядка нескольких дней) элементы орбиты могут вообще не обновляться. Для расчёта коэффициента корреляции необходимо сформировать пары параметров  $V^*$ , отнесённые к одному и тому же моменту времени.

Для решения этой задачи для временного ряда каждого спутника отдельно вычислялись дистанции во времени между двумя соседними оценками элементов орбиты КО. Затем для сетки значений промежутков времени 0, 0.01, 0.02, ..., 5 суток вычислялась вероятность, что дистанция между соседними оценками параметров орбиты меньше заданного. Таким образом, задавая вероятность иметь дистанцию меньше заданной можно получить оценку времени дискретизации временной шкалы для каждого КО отдельно. Далее для каждой пары объектов выбиралась наибольшая оценка времени дискретизации и вся шкала времени дискретизировалась с выбранным шагом для обоих спутников. Внутри каждого дискретного интервала значения параметра  $V^*$  усреднялись. Если для хотя бы одного спутника на дискретном интервале отсутствовали оценки параметра  $V^*$ , то пара отбрасывалась. Для приведенных ниже примеров, вероятность иметь дистанцию меньше заданной, выбиралась равной 0,5. По нашему мнению, это компромиссное значение между излишней дискретизацией временной шкалы и излишним заглаживанием колебаний вариаций торможения спутника в результате усреднения параметра  $V^*$ .

Теперь коэффициенты корреляции легко могут быть рассчитаны по хорошо известной формуле

$$\rho_{jk} = \frac{\sum_i (B_{ij} - \bar{B}_j)(B_{ik} - \bar{B}_k)}{\sqrt{\sum_i (B_{ij} - \bar{B}_j)^2 \sum_i (B_{ik} - \bar{B}_k)^2}}, \quad (1)$$

где  $\bar{B}_j = \frac{\sum_{i=1}^N B_{ij}}{N}$  – среднее значение параметра  $V^*$ .

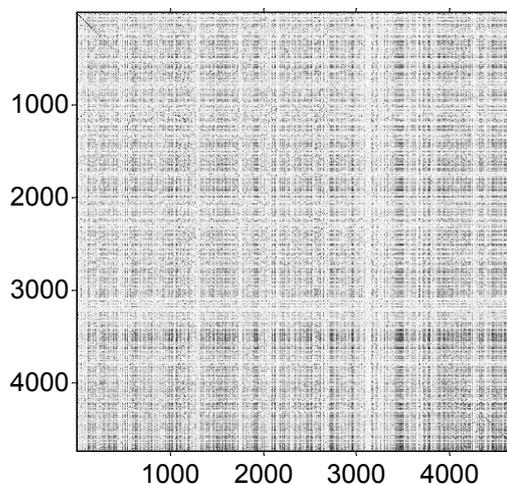


Рис. 3. Корреляционная матрица параметров  $V^*$  для всех объектов (для 2007 года). Чёрным цветом показаны коэффициенты корреляции большие 0,5.

Как видно (рис. 3) существует достаточно большое множество объектов, у которых наблюдаются похожие изменения параметра  $V^*$  со временем. Для формирования групп КО похожих в этом смысле друг на друга применим кластерный анализ.

**Кластерный анализ изменения торможения ИСЗ**

Для реализации кластерного анализа необходимо ввести пространство, метрику и расстояние в пространстве объектов. Следует ожидать, что объекты, имеющие сходные изменения параметра  $V^*$  со временем, будут иметь близкие значения коэффициентов корреляции со всеми объектами. Иначе говоря, строки корреляционной матрицы на рис. 3 будут похожи друг на друга. В этом случае норма разности между строками будет характеризовать расстояние между объектами. Однако выбор базиса из объектов представляется не лучшим в виду его неортогональности и неортонормированности. Кроме того плотность вероятности коэффициента корреляции заметно отличается от плотности вероятности нормального распределения. Поэтому введём величину  $y$  определяемую как

$$y = \frac{1}{2} \ln \left( \frac{1+\rho}{1-\rho} \right). \tag{2}$$

Плотность вероятности величины  $y$  совпадает с плотностью вероятности нормального распределения [4].

Для расчёта ортогонального ортонормированного базиса рассчитаем спектр собственных значений матрицы  $y_{jk}$  (рис. 4). На рисунке приведены первые 110 наибольших по абсолютному значению собственных чисел.

Спектр собственных значений быстро переходит к шумовому характеру. Для большинства прикладных задач можно ограничиться двадцатью первыми собственными значениями. Однако сейчас мы возьмём 100 собственных значений и для них рассчитаем собственные вектора, которые создадут 100 мерный ортогональный ортонормированный базис в пространстве коэффициентов корреляции.

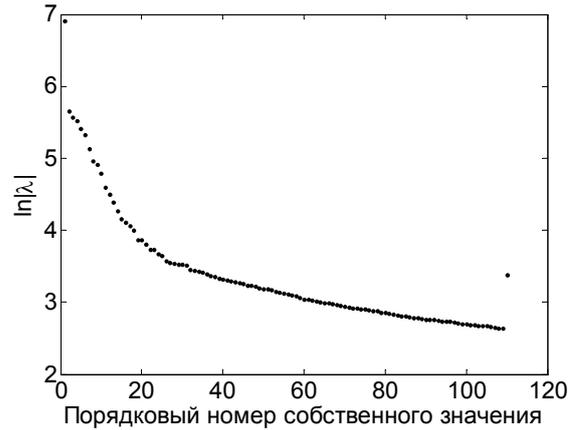


Рис. 4. Спектр 110 наибольших по абсолютному значению собственных чисел корреляционной матрицы  $y_{jk}$ .

Тогда координаты каждого объекта в пространстве собственных векторов определяются как

$$C_{jq} = \sum_k y_{jk} V_{kq}, \tag{3}$$

где  $V_{kq}$  – собственные вектора.

Для формирования кластеров воспользуемся методом Варда [5]. На рисунке 5 представлена дендрограмма распределения 30 кластеров. Как видно дерево кластеров хорошо сбалансировано. Сделаем теперь переупорядочивание объектов в соответствии с кластерным анализом, в результате которого корреляционная матрица принимает новый вид (рис. 6). Хорошо просматривается лучшая структурированность матрицы, чем на рисунке 3. Это свидетельствует о правильности формирования кластеров. Следует ожидать, что у объектов принадлежащих одному кластеру будет наблюдаться большее сходство в изменениях параметров  $V^*$  от времени, чем у объектов принадлежащих двум разным кластерам.

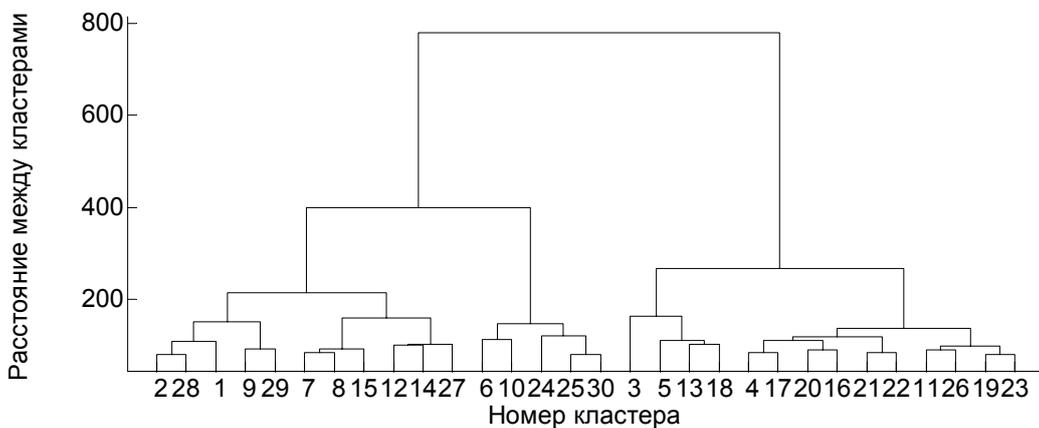


Рис. 5. Дендрограмма распределения и родственности кластеров.

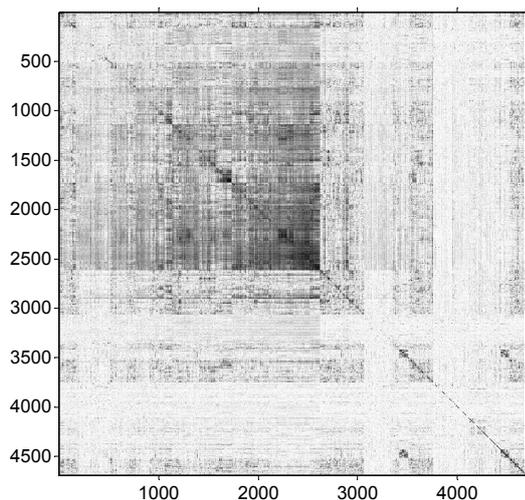


Рис. 6. Корреляционная матрица перепорядоченная в соответствии с кластерным анализом и дендрограммой на рис. 5.

Наконец продемонстрируем распределение количества КО в разных кластерах при общем количестве кластеров равном 200 (рис. 7).

Проанализируем распределение орбитальных параметров КА, входящих в один кластер. Как правило это слабо эллиптические орбиты. Однако, во многих кластерах присутствуют и заметно эксцентричные орбиты. Для наглядного сопоставления между собой формы орбит в разных кластерах отсортируем высоты перигеев орбит внутри кластера и будем последовательно располагать номера КО во всех кластерах по оси абсцисс (рис. 8). В этом случае резкое снижение высоты перигея служит точкой разграничивающей кластеры. Для простоты рассмотрим разбиение на 30 кластеров. Видно, что почти все кластеры включают орбиты с широким набором высот перигея – от минимальной до 1000 км.

Если на этом же графике изобразить соответствующие высоты апогеев орбит, то получим представление и о форме орбит. Некоторые большие кластеры явно включают орбиты разного свойства, то есть, возможно, являются эклектичными и требуется их дальнейшее деление на классы.

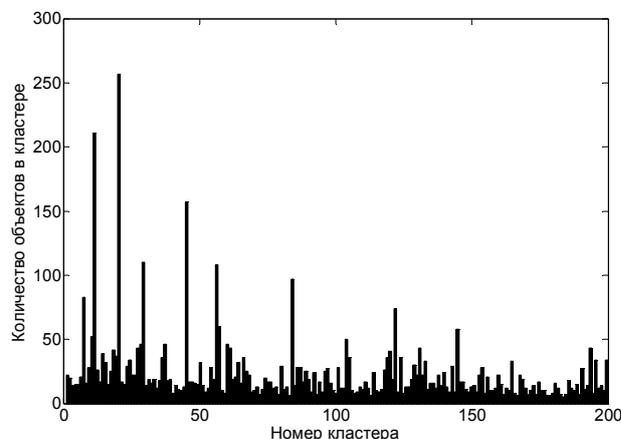


Рис. 7. Распределение количества объектов от номера кластера.

Для того чтобы оценить пользу от проделанной работы по кластеризации рассмотрим несколько видов поведения  $V^*$  со временем для больших кластеров.

Разные КО в кластере демонстрируют синхронное изменение  $V^*$  со временем, при этом наблюдаются как быстрые колебания так и их низкочастотная модуляция разного характера. Это может быть отражением глобальных колебательных изменений плотности верхней атмосферы. Однако, прежде эти эффекты необходимо отделить от периодического изменения параметров самой орбиты КА и ее ориентации относительно приливных и термических структур в атмосфере.

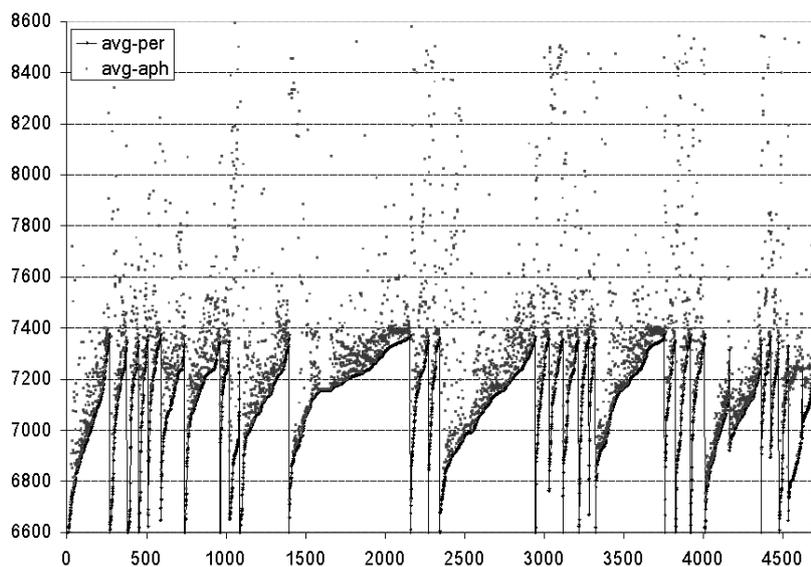


Рис. 8. Распределение высот перигеев и апогеев в 30 кластерах.

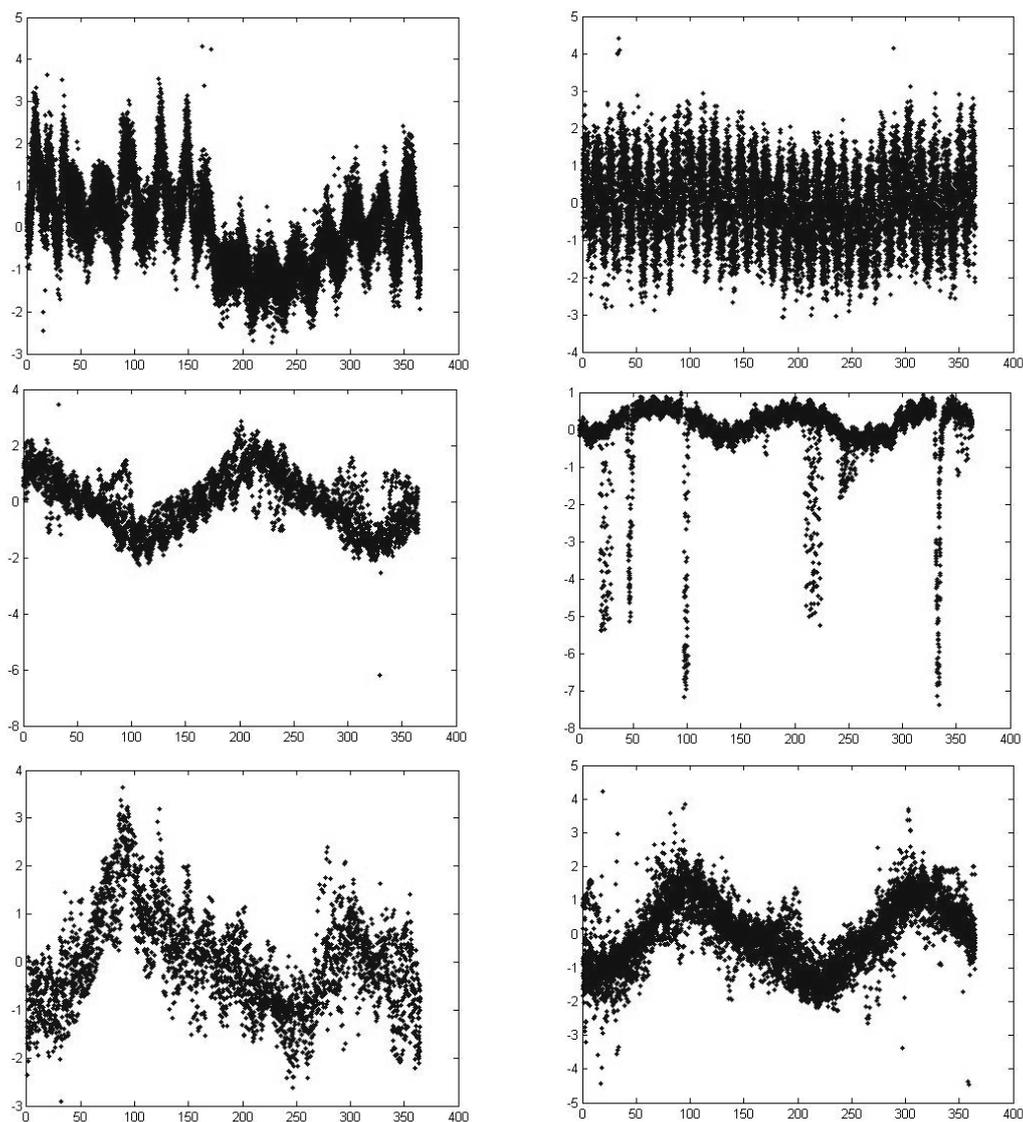


Рис. 9. Наложение графиков изменения  $V^*$  со временем для всех КО в отдельных кластерах.

### Заключение

На основе собственных наблюдательных данных и доступных каталогов орбит ИСЗ сформирована значительная база орбитальных данных, которая может быть полезна для развития теории движения ИСЗ под влиянием сопротивления атмосферы. В результате проделанного корреляционного и кластерного анализа характера изменения баллистического параметра  $V^*$  со временем у более чем четырех тысяч космических объектов на низких околоземных орбитах выявлены большие группы КО, с подобным характером изменения этой величины внутри группы и существенно различным видом этого изменения от группы к группе. Этот факт требует дальнейшего осмысления и анализа на новом наблюдательном материале. Отметим, что сделанное нами сопоставление вариаций плотности атмосферы, полученных на основе измерений торможения для ряда КО на разных орбитах, и “теоретических” вариаций плотности вдоль тех же орбит полученных для модели MSISE-2000, показывает значительные

отличия и снижает доверие к модели. С другой стороны, мы не нашли заметной корреляции между реальными вариациями скорости торможения ряда КО и циклическим изменением положения их орбит по отношению к Солнцу и Луне, как источникам приливных эффектов в атмосфере.

### Литература

1. <http://www.celestrak.com/NORAD/documentation/tle-fmt.asp>
2. <http://www.celestrak.com/publications/AIAA/20066753/AIAA-2006-6753.pdf>
3. N.Koshkin at all. *Sun and Geosphere*. – 2006. – Vol.1(2). – P. 46–49. ([http://www.shao.az/SG/v1n2/SG\\_v1\\_No2\\_2006-pp-46-49.pdf](http://www.shao.az/SG/v1n2/SG_v1_No2_2006-pp-46-49.pdf))
4. Г.Корн и Т.Корн. Справочник по математике для научных работников и инженеров. М., Наука, 1978. 832 с.
5. Ward J. H. Hierarchical grouping to optimize an objective function. *J. Am. Statist. Assoc.*, 1963, vol. 58, pp. 236-244.