

НЕКОТОРЫЕ РЕЗУЛЬТАТЫ СТАТИСТИЧЕСКОГО АНАЛИЗА ОПРЕДЕЛЕНИЙ ГАЛАКТОЦЕНТРИЧЕСКОГО РАССТОЯНИЯ СОЛНЦА

Малкин З.М.^{1,2}

¹Главная (Пулковская) астрономическая обсерватория РАН, Санкт-Петербург, Россия

²Астрономический институт им. В.В. Соболева СПбГУ, Санкт-Петербург, Россия

*К настоящему времени накопилось несколько десятков определений расстояния от Солнца до центра галактики R_0 , полученных разными методами и по разному набору объектов. В работе проведен статистический анализ 52 результатов, опубликованных за последние 20 лет. Основными целями этого исследования были изучение статистической согласованности этих разнородных данных, определение наилучшего среднего значения и проверка наличия в данных "эффекта присоединения к большинству" (ЭПБ, *bandwagon effect*), подозреваемого некоторыми авторами. Оказалось, что результаты определения R_0 являются статистически согласованными. Наилучшее среднее значение, вычисленное по совокупности нескольких статистических методов, получилось равным 8.0 ± 0.25 кпс. Статистически значимого тренда со временем в этих данных не обнаружено, что не подтверждает выводы некоторых предыдущих работ. Отсутствие тренда также ставит под сомнение наличие в результатах определения R_0 ЭПБ. В то же время найдено статистически значимое уменьшение ошибок определения R_0 со временем.*

Введение

История этих определения расстояния от Солнца до центра Галактики (R_0) насчитывает уже около ста лет, в течение которых были сделаны несколько десятков измерений R_0 разными методами, по разным наборам объектов, и с разной точностью. В то же время астрономическое сообщество нуждается в некотором стандартном значении R_0 , соответствующем современным определениям. В частности, побудительным мотивом для настоящей работы послужила необходимость улучшения точности вычисления аберрации в собственных движениях (Галактической аберрации) внегалактических радиоисточников [2]. За прошедшие годы было предпринято несколько попыток получить среднее значение из имеющихся данных. Как правило, опубликованные работы основаны на тщательном звездно-астрономическом анализе имеющихся результатов, но ни в одной не сделана попытка их серьезного статистического анализа. Вместе с тем, как показывает опыт определения фундаментальных констант, например, в физике, эта задача является не только специально-научной, но и метрологической. В настоящей работе проведен анализ 52 определений R_0 , сделанных в течение последних 20 лет разными статистическими методами. Основные целями исследования были вычисление наилучшего среднего значения и его реалистичной ошибки, а также определение тренда в данных, который может свидетельствовать об эффекте присоединения к большинству (ЭПБ, в англоязычной литературе *bandwagon effect*) в опубликованных значениях R_0 , что иногда обсуждается в литературе.

Детали этой работы могут быть найдены в [7].

Использованные данные

Одной из причин расхождения результатов предыдущих авторов может быть селекция использованных данных в результате их анализа, предпочтений автора или из-за недоступности публикаций. Чтобы избежать этого, мы постарались собрать все данные, опубликованные за последние 20 лет, за исключением трех, пересмотренных авторами позднее. В результате, было использовано 52 значения R_0 (табл. 1).

Таблица 1. Используемые результаты определения R_0 .

7.9 ± 0.8	Merrifield, M. R. 1992, AJ, 103, 1552
8.1 ± 1.1	Gwinn, C. R., et al. 1992, ApJ, 393, 149
7.6 ± 0.6	Moran, J. M., et al. 1993, Lecture Notes in Physics, 412, 244
7.6 ± 0.4	Maciel, W. J. 1993, Astrophys. and Space Sci., 206, 285
8.09 ± 0.3	Pont, F., et al. 1994, Astron. Astrophys., 285, 415
7.5 ± 1.0	Nikiforov, I. I., Petrovskaya, I. V. 1994, Astron. Rep., 38, 642
7.0 ± 0.5	Rastorguev, A. S., et al. 1994, Astron. Lett., 20, 591
8.8 ± 0.5	Glass, I. S., et al. 1995, MNRAS, 273, 383
7.1 ± 0.5	Dambis, A. K., et al. 1995, Astron. Lett., 21, 291
8.3 ± 1.0	Carney, B. W., et al. 1995, AJ, 110, 1674
8.21 ± 0.98	Huterer, D., et al. 1995, AJ, 110, 2705
7.95 ± 0.4	Layden, A. C., et al. 1996, AJ, 112, 2110
7.55 ± 0.7	Honma, M., Sofue, Y. 1996, PASJ, 48, L103
8.1 ± 0.4	Feast, M. W. 1997, MNRAS, 284, 761
8.5 ± 0.5	Feast, M., Whitelock, P. 1997, MNRAS, 291, 683
7.66 ± 0.54	Metzger, M. R., et al. 1998, AJ, 115, 635
8.1 ± 0.15	Udalski, A. 1998, Acta Astron., 48, 113
7.1 ± 0.4	Olling, R. P., Merrifield, M. R. 1998, MNRAS, 297, 943
8.51 ± 0.29	Feast, M., et al. 1998, MNRAS, 298, L43
8.2 ± 0.21	Stanek, K. Z., Garnavich, P. M. 1998, ApJ, 503, L131
8.6 ± 1.0	Surdin, V. G. 1999, Astron. Astrophys. Trans., 18, 367
7.4 ± 0.3	Glushkova, E. V., et al. 1999, Astron. Astrophys. Trans., 18, 349
7.9 ± 0.3	McNamara, D. H. et al. 2000, PASP, 112, 202
8.67 ± 0.4	Stanek, K. Z., et al. 2000, Acta Astron., 50, 191
8.2 ± 0.7	Nikiforov, I. I. 2000, ASP Conf. Ser., 209, 403
8.24 ± 0.42	Alves, D. R. 2000, ApJ, 539, 732
8.05 ± 0.6	Genzel, R., et al. 2000, MNRAS, 317, 348
8.3 ± 0.3	Gerasimenko, T. P. 2004, Astron. Rep., 48, 103
7.7 ± 0.15	Babusiaux, C., Gilmore, G. 2005, MNRAS, 358, 1309
8.01 ± 0.44	Avedisova, V. S. 2005, Astron. Rep., 49, 435
7.62 ± 0.32	Eisenhauer, et al., 2005, ApJ, 628, 246
8.7 ± 0.6	Groenewegen, M. A. T., Blommaert, J. A. D. L. 2005, Astron. Astrophys., 443, 143
7.2 ± 0.3	Bica, E., et al. 2006, Astron. Astrophys., 450, 105
7.52 ± 0.36	Nishiyama, S., et al. 2006, ApJ, 647, 1093
8.1 ± 0.7	Shen, M., Zhu, Z. 2007, Chin. J. Astron. Astrophys., 7, 120
7.4 ± 0.3	Bobylev, V. V., et al. 2007, Astron. Lett., 33, 720
7.94 ± 0.45	Groenewegen, M. A. T., et al. 2008, Astron. Astrophys., 481, 441
8.07 ± 0.35	Trippe, S., et al. 2008, Astron. Astrophys., 492, 419
8.16 ± 0.5	Ghez, A. M., et al. 2008, ApJ, 689, 1044
8.33 ± 0.35	Gillessen, S., et al. 2009, ApJ, 692, 1075
8.7 ± 0.5	Vanhollebeke, E., 2009, Astron. Astrophys., 498, 95
7.58 ± 0.40	Dambis, A. K. 2009, MNRAS, 396, 553
7.2 ± 0.3	Bonatto, C., et al. 2009, Globular Clusters---Guides to Galaxies, 209
8.4 ± 0.6	Reid, M. J., et al. 2009, ApJ, 700, 137
7.75 ± 0.5	Majaess, D. J., et al. 2009, MNRAS, 398, 263
7.9 ± 0.75	Reid, M. J., et al. 2009, ApJ, 705, 1548
8.24 ± 0.43	Matsunaga, N., et al. 2009, MNRAS, 399, 1709
8.28 ± 0.33	Gillessen, S., 2009, ApJ, 707, L114
7.7 ± 0.4	Dambis, A. K. 2010, Variable Stars, the Galactic halo and Galaxy Formation, 177
8.1 ± 0.6	Majaess, D. 2010, Acta Astron., 60, 55
8.3 ± 1.1	Sato, M., et al. 2010, ApJ, 720, 1055
7.80 ± 0.26	Ando, K., et al. 2011, PASJ, 63, 45

В тех случаях, когда авторы приводили оценки и случайных и систематических ошибок, эти ошибки складывались в квадратуре. Иногда авторы приводили в качестве оценки точности не среднеквадратическую ошибку, а несимметричный доверительный интервал. В таких случаях в качестве ошибки результата принималось среднее из верхней и нижней границы этого интервала. Поскольку эти границы во всех случаях оказались близки между собой, такая замена не оказала заметного влияния на полученные в настоящей работе результаты.

Анализ данных

Как уже сказано выше, анализ данных проводился в двух основных направлениях. Первой задачей было определение наилучшей оценки R_0 с учетом последних наблюдений. Не менее важным для дальнейшего использования было получить реалистичную оценку ошибки этого значения. Для решения задачи мы использовали несколько статистических методов, применяющихся при определении наилучших значений физических констант и описанных в [7] и в цитированной там литературе. Шесть из них являются вариантами метода наименьших квадратов (МНК), а два метода основаны на анализе медианных значений. Эти методы следующие:

1. МНК без весов;
2. МНК с весами, обратно пропорциональными ошибкам;
3. МНК с весами и модифицированным методом вычисления ошибки [3];
4. Метод ограничения относительных весов [9];
5. Метод нормализованных уклонений [5];
6. Метод Мандела-Пауле [11];
7. Медиана с вычислением ошибки результата по [8];
8. Бутстреп медиана.

Методы 4–6 являются вариантами МНК с весами, отличающиеся тем, что в них разными способами выделяются и перевзвешиваются исходные значения, статистически не согласующиеся с остальными. Результаты вычисления среднего значения R_0 каждым методом приведены в табл. 2.

Таблица 2. Результаты вычисления среднего значения R_0 разными методами.

Метод	R_0 , кпс
Невзвешенное среднее	7.960 ± 0.062
Взвешенное среднее	7.909 ± 0.051
Взвешенное среднее (модифицированное)	7.909 ± 0.076
Ограничение относительных весов	7.909 ± 0.051
Нормализованные уклонения	7.909 ± 0.051
Мандел-Пауле	7.911 ± 0.051
Медиана	8.060 ± 0.075
Бутстреп медиана	8.028 ± 0.083

Прежде всего, следует заметить, что средние значения R_0 , полученные разными модификациями МНК с весами (методы 2–5), практически одинаковы, что говорит о статистической согласованности данных, хотя ошибки их различаются иногда существенно. В то же время, разброс средних значений, полученными разными методами, в некоторых случаях превышает стандартные ошибки, причем нет никаких теоретических или практических причин отдать предпочтение одному из них перед остальными. Надо заметить, что такая ситуация является типичной в такого рода задачах.

При таких обстоятельствах наилучшее среднее значение может быть получено из комбинации приведенных результатов. Мы вывели наилучшую оценку R_0 следующим

образом. Окончательное значение R_0 взято как середина интервала, покрываемого всеми полученными средними с их стандартными ошибками. В нашем случае это среднее между крайними значениями ($7.909-0.076$) и ($8.060+0.075$), равное 7.984 . Ошибка наилучшего значения может быть принята как половина этого интервала, т.е. 0.151 . Однако это только статистическая оценка ошибки. Хотя многие авторы рассматривали систематические ошибки определений R_0 разными методами, но их точная величина, разумеется, неизвестна (иначе она была бы учтена в публикуемых результатах!). Сопоставляя различные литературные данные и учитывая, что наше значение R_0 получено усреднением данных, полученных различными методами (что должно приводить к компенсации возможных систематических ошибок этих методов), мы считаем разумным пределом среднеквадратической величины систематической ошибки окончательного значения 0.2 кпс. Комбинируя все эти данные, мы получили окончательное значение $R_0=8.0\pm 0.25$, которое считаем наилучшей оценкой галактоцентрического расстояния Солнца на настоящий момент. Наше решение вместе с другими сводными значениями R_0 приведено в табл. 3. Можно видеть, что наилучшая оценка R_0 за последние 20 лет практически не изменилась.

Таблица 3. Все результаты вычисления среднего значения R_0 .

Работа	Интервал данных	Число измерений	R_0 , кпс
[6]	1974–1986	25	8.5 ± 1.1
[12]	1974–1987	30	7.7 ± 0.7
[13]	1974–1992	35	8.0 ± 0.5
[10]	1974–2003	65	7.9 ± 0.2
[1]	1992–2005	16	7.8 ± 0.32
Наст. работа	1992–2011	52	8.0 ± 0.25

Следующим этапом работы стало изучение тренда в изменении опубликованных значений R_0 со временем. На наличие тренда обращали внимание многие авторы, однако часто они приходили к разным выводам в отношении его величины и даже направления. Так в данных [10, 13, 14] присутствует значительный отрицательный наклон в период с 1970-х до начала 1990-х гг. (этот интервал в настоящей работе не рассматривается), в данных [14] практически отсутствует тренд в 1990–1998 гг., в данных [10] присутствует небольшой положительный тренд в 1990–2003 гг., а в [4] показан большой положительный тренд в 1992–2010 гг.

Для наших вычислений мы использовали данные табл. 1, которые показаны на рис. 1. По ним был определен линейный тренд МНК в двух вариантах: без использования весов и с весами, обратно пропорциональными ошибкам. В результате получилось значение тренда -0.011 ± 0.011 кпс/год для вычислений с весами и $+0.006\pm 0.010$ кпс/год для вычислений без весов. Таким образом, мы можем заключить, что статистически значимого тренда в определениях R_0 за последние 20 лет не обнаружено.

Этот вывод подкрепляется также результатами проверки наличия систематических изменений в данных по критерию Аббе. Для этого вычисляется статистика Аббе q , которая может быть представлена как отношение вариации Аллана $AVAR$ к несмещенной оценке дисперсии выборки D :

$$q = \frac{AVAR}{D}, \quad AVAR = \frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{2(n-1)}, \quad D = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}. \quad (1)$$

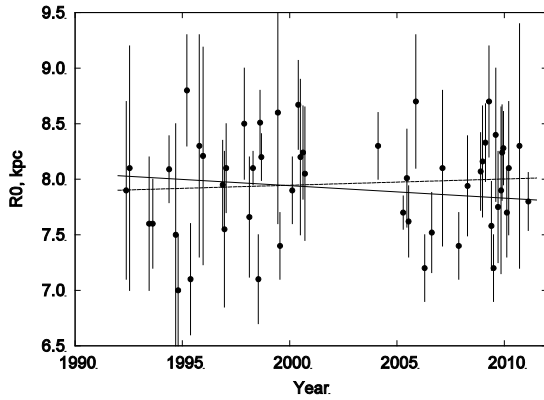


Рис. 1. Используемые определения R_0 и линейный тренд (сплошная линия – взвешенный, пунктирная линия – невзвешенный), кпс, кпс/год.

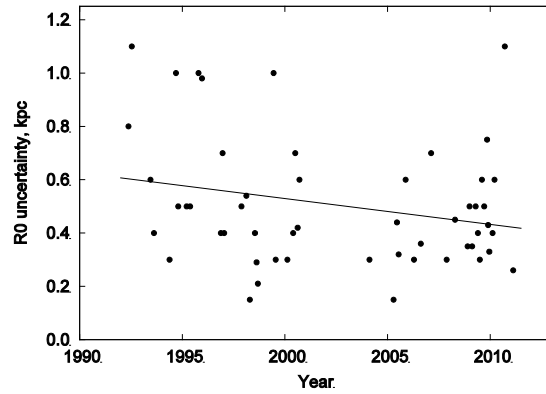


Рис. 2. Ошибки определения R_0 и линейный тренд, кпс, кпс/год.

В присутствии значительных систематических (низкочастотных) вариаций в данных величина D будет существенно больше $AVAR$, а q , соответственно, существенно меньше единицы. В нашем случае оказалось $q = 1.26$, что почти вдвое больше 1% критического значения 0.69. Впрочем, любое значение $q > 1$ уже говорит об отсутствии в данных статистически значимых низкочастотных изменений и тренда.

Причины расхождения нашего вывода об отсутствии тренда с результатом [4] видны из сравнения рис. 1 здесь с рис. 1 в работе [4]. Авторы последней по необъясненной причине не использованы несколько значений R_0 больше среднего из первой половины интервала и несколько значений меньше среднего из второй половины интервала, что и привело к найденному ими значительному росту публикуемых значений R_0 со временем. Дальнейший анализ показывает, что и расхождения между другими работами в отношении трендов объясняется разными наборами исходных данных.

Последней частью нашей работы было изучение изменения со временем ошибки определения R_0 . Соответствующие данные приведены на рис. 2 вместе с линией, соответствующей найденному МНК линейному тренду -0.0097 ± 0.0054 кпс/год. Таким образом, точность определения R_0 существенно улучшилась за последние 20 лет. Надо заметить, что в последние годы авторы стали обращать больше внимания на получение реалистичных ошибок своих результатов. Поэтому, с одной стороны, ошибки R_0 должны со временем уменьшаться по мере накопления и улучшения наблюдательных данных и используемых методов, а с другой стороны, эти ошибки могут увеличиваться из-за использования более корректных методов их оценки. Очевидно, первая тенденция в последние годы оказалась сильнее.

Заключение

В работе произведен статистический анализ 52 определений галактоцентрического расстояния Солнца R_0 , сделанных за последние 20 лет. При этом использовались несколько методов, что позволило лучше исследовать статистические свойства этих оценок. Хотя измерения были сделаны разными методами и по разным наборам объектов, они оказались согласованными в статистическом смысле. Заметим, что этот вывод не является тривиальным. Предварительный анализ определений угловой скорости вращения Галактики Ω_0 , сделанных за тот же период, показал, что эти измерения статистически не согласованы.

В качестве одного из основных результатов работы было вычислено среднее значение $R_0 = 8.0 \pm 0.25$, что находится в хорошем согласии с последними определениями $R_0 = 8.05 \pm 0.45$ [15]. Было также изучено изменение опубликованных значений R_0 со

временем. Статистически значимого тренда в этих данных не обнаружено, что может означать отсутствие в них ЭПБ, возможное наличие которого обсуждается несколькими авторами, например, [10,13]. В то же время, тренд в опубликованных ошибках определения R_0 получился равным $-0:0097 \pm 0054$ кпс/год, что говорит о статистически значимом улучшении точности измерений со временем.

Литература

1. *Аведисова В.С.* Галактические постоянные и кривая вращения Галактики по молекулярному газу. *Астрон. журн.*, 2005, т. 49, 488–499.
2. *Малкин З.М.* Влияние галактической аберрации на параметры прецессии, определяемые из РСДБ-наблюдений. *Астрон. журн.*, 2011, т. 88, 880–885.
3. *Малкин З.М.* О вычислении ошибки среднего взвешенного. *Наст. выпуск.*
4. *Foster T., Cooper B.* Structure and Dynamics of the Milky Way: The Evolving Picture. In: ASP Conf. Ser., v. 438, 2010, 16–30.
5. *James M.F., Mills R.W., Weaver D.R.* Nucl. Instr. Meth. Phys. Res. A, 1992, v. 313, 277–282.
6. *Kerr F.J., Lynden-Bell D.* Review of galactic constants. *MNRAS*, 1986, v. 221, 1023–1038.
7. *Malkin Z.* The current best estimate of the Galactocentric distance of the Sun based on comparison of different statistical techniques. arXiv:1202.6128, 2012.
8. *Müller J.W.* Possible advantages of a robust evaluation of comparisons. Report BIPM-95/2. BIPM, Sevres, France, 1995.
9. *Nichols A.L.* IAEA Co-ordinated Research Project: update of X-ray and gamma-ray decay data standards for detector calibration and other applications. *Appl. Rad. Isot.*, 2004, v. 60, 247–256.
10. *Nikiforov I.* The Distance to the Center of the Galaxy: the Current State-of-the-Art in Measuring R_0 . In: ASP Conf. Ser., v. 316, 2004, 199–208.
11. *Paule R., Mandel J.* Consensus Values and Weighting Factors. *J. Res. Natl. Bur. Stand.*, 1982, v. 87, 377–385.
12. *Reid M.J.* The Distance to the Galactic Center: R_0 (review). In: IAU Symposium, v. 136, The Center of the Galaxy, 1989, 37–46.
13. *Reid M.J.* The distance to the center of the Galaxy. *ARA&A*, 1993, v. 31, 345–372.
14. *Surdin V.G.* Distance to the galactic centre. *Astron. Astrophys. Trans.*, 1999, v. 18, 367–372.
15. *Honma M., Nagayama T., Ando K., et al.* Fundamental Parameters of the Milky Way Galaxy Based on VLBI astrometry. arXiv:1211.3843v2, 2012

SOME RESULTS OF STATISTICAL ANALYSIS OF THE DETERMINATIONS OF THE GALACTOCENTRIC DISTANCE OF THE SUN

Malkin Z.M.^{1,2}

¹*Central Astronomical Observatory at Pulkovo of RAS, St. Petersburg, Russia*

²*Sobolev Astronomical Institute, St. Petersburg State University, St. Petersburg, Russia*

Based on R_0 measurements made during last decades, several studies were performed to derive the best R_0 estimate. Some of them used just simple averaging to get a result, whereas others provided comprehensive analysis of possible errors in published results. In both cases no detailed statistical analysis of data used was made. However, a computation of the best estimates of the Galactic rotation constants is not only astronomical but also metrological task. In this work we performed an analysis of 52 R_0 measurements published during last 20 years to assess the consistency of the data. Our analysis has shown that they are consistent (unlike Ω_0). The average value of R_0 found in our study is 8.0 ± 0.25 kpc. It was also shown that the trend in the R_0 estimates for the last 20 years is statistically negligible, which makes it doubtful a presence of the bandwagon effect. On the other hand, the formal error of the published R_0 estimates significantly improves with time.