

MetroWest Chess Club Player Distribution Analysis

Author: Mark Kaprielian

2014-07-04

Table of Contents

Author: Mark Kaprielian.....	1
I. Purpose.....	2
II. Attendance in relation to Club Location.....	2
III. The Past and Present.....	3
IV. Viewing the Modern Era of the club via the numbers.....	4
V. Keeping a current eye on the numbers.....	5
VI. Some new views of our data.....	6
VII. The new reports.....	7
A. Section Averages.....	8
B. Balance Between sections.....	10
C. Balance Between sections as a percentage.....	11
VIII. Modern 80 and standard breaks.....	12
A. Open Section.....	12
B. Open Section – 3 Levels.....	13
C. U2000 Section.....	14
D. U1700 Section.....	15
E. U1400 Section.....	16
IX. All Years.....	17
A. Open Section participation.....	17
B. Open Section - 3 Levels.....	18
C. Open Section – 2 Levels.....	19
D. Open Section – 1 Levels.....	20
X. Modern 80 era.....	21
A. Open section.....	21
B. Open section – 3 Levels.....	22
XI. Answering the Questions.....	23
A. What are the most masters we’ve had play in an event?.....	23
B. What would be a reasonable definition of “strong Open section”?.....	23
1. Average of ratings in the Open section.....	23
2. Ratio of Masters, Experts and Class players in the Open section.....	23
3. Count of Masters, Experts and Class players in the Open section.....	23
C. What is the strongest Open section we’ve had?.....	24
D. Summary of the answers.....	24

End of Table of Contents

I. Purpose

The purpose of this document is to provide some insight into how we learn from our experiences running the club. In this document I will provide some background of how we view our past events and how we may use these views to guide our decisions for the future.

Participation and attendance are hard, possibly impossible things to predict yet, we have the desire and need to plan so we try, however fruitlessly to predict the future based on the past. Cross tables provide the historical record of what was. By examining our cross tables and trying to correlate our decisions with the historical record, we hope to be able to learn and thus control our fate.

We have had at our disposal a database that was created by us to contain the crossable of every event the club has ever held. Each month we append in the previously completed event's cross table and produce a collection of reports that we publish to the club's web site.

In addition to the regular reports we produce we often do specialized examinations of the data to try and root-out patterns and useful information. For example a while back we created the ability to export out to excel and graph how many players are playing up or playing in their "natural" section. We gained some valuable insight from this data. Often this kind of data isn't easy or pretty to present so we don't necessarily post it to the web site as it isn't in a generally consumable form.

Some of the more visible standard "data crunching" activities that occur on a regular basis are for example:

- Generating the list of who should be awarded their next participation award button.
- Another appears on your pairing label sticker which you may have noticed and wondered about forming this question in your mind, "what does plago mean and why is it taking up space on this tiny sticker? " I'm not telling here, you'll have to ask a TD if you haven't already.

II. Attendance in relation to Club Location

Starting in 1995 when I was elected president of the club, attendance had been dropping off and the primary volunteers were exhausted. My first night as president was also the first time the club had meet in a venue other than its original home of 13 years. While we had a spike in interest for the first month, attendance quickly returned to its low end and many players were very unimpressed with the new location which was "a pit". We played between piles of old furniture and other "stuff". We could barely accommodate 20 players. It took us a year to find another location, the Natick Community center, which prior to it being rebuilt in 2013 was called the Natick Senior Center, where we have been ever since. The Natick Senior center was an old style elementary school that had been repurposed. This provided us with the ability to fit, after we purchased the tables for the Senior center, 80 players comfortably in one main room and a skittles room where we could accommodate another 24 players when needed. With a new location that was clean and "spacious" according to many players who had played in a wide range of clubs both locally and out of state, we were able to start promoting the club without fear of running out of space or having a location that people would not want to return to. When we moved to this location the club was at an all-time low for attendance of about 20 players and in the previous few years had only a few occasions reaching 32 players.

With the move from Framingham to Natick we risked losing a few players from the West but we now become more enticing to players from the East. A demographic study that I had performed using USCF address data available to plot physical distribution of active tournament players in Massachusetts had the city of Newton Mass as its center. This is not to say that Newton was where the most players were but that if you selected a weighted epicenter of physical location, it was centered in Newton. Basically this was due to players out in the MetroWest Area balanced by those up to the north on Rt. 128 such as the Waltham and Billerica chess clubs and the Boylston Chess Club which was the center for all playing in Boston and Cambridge.

III. The Past and Present

As of this writing the club is in its 31st year of existence. Since it's founding in 1983 there have been several milestone events.

- 1983: the club is started and run primarily by its founder, Warren Pinches
- 1988: Warren Pinches, the driving force of the club since its start, retires from ongoing operations of the club.
- 1988 – 1995: The club while run by a dedicated and hard working group of people begins to experience drop off in both volunteers and attendance
- 1995: The club is forced to leave its original home of 12 years due to rising rent and dwindling income. The club relocates in the basement of the Framingham Senior center with a very limited space and lots of piled up old furniture. Access to the building depends on the door being jammed open by a door stop. Attendance drops to 12 to 22 players.
- 1996: The club relocates to the Natick Senior center.
- 1997: The club web site goes up and has weekly updates of results and presents projected pairings. Note: This is likely the first time this had ever been done by a weekly club.

With such a long history and major changes in location and technology used by the club it is useful to think of there being several “era’s”.

- **“All Years”** – From the clubs inception in 1983 to the present
- **“Early years”** or **“Pre-Modern”** – Characterized as being the year’s most strongly influenced by the guidance of Warren Pinches and his legacy till about 1994
- **“Transitional”** – Characterized as starting with the 1st move of the clubs location in 1995 and ending around the time of the 2nd move to its location in Natick in 1996
- **“Modern”** – Characterized as starting with the introduction of the club’s website in 1997
 - **“Modern 80”** – This represents the Modern Era years where our average fist went above 80 players, 2003.
 - **“Modern 80 and standard breaks”** - This represents the Modern Era years where our average went to 80 players and we fixed our sections at Open, U2000, U1700, U1400. This starts in July of 2009

IV. Viewing the Modern Era of the club via the numbers

The move to Natick started a series of growth spurts and plateaus. This can be observed by the graph that is updated monthly on the club web site, see the link “How many Play ??”.

http://www.metrowestchess.org/Community/People/Demographics/Attendance/MCC_Chart_of_Turnout_Monthly.pdf

For years we have revisited the seasonal attendance and have occasionally observed some impact from the weather and some historical events (Boston teams in play-offs) but no strong correlations have been agreed upon though there are some minor trends.

On the “How Many Play ??” graph linked from the web site home page we present the Monty attendance via data points. As you will observe the attendance can vary widely from one month to the next. One metric that has proven useful is the average attendance for the calendar year, as displayed by the pink line on the graph.

Using this graph we can provide a profile of the events leading to the growth of the club attendance.

- Prior to 1997 a lot of fluctuation between 23 and 32 players (yearly average).
- In 1997 after settling into our current location in Natick the first half of the year we seemed to be steady at the higher end of attendance numbers that we'd seen in the past five years and in the second half of the year we began to see a clustering of higher attendance numbers than we had seen for a long time. The result of this was that the average for 1997 was higher than the preceding five years.
- In 1997 we also introduced the club's web site and introduced weekly results and projected pairings. Players, both those in the event and those in with an interest in local chess could see who was playing and what the results were the next day. It is my belief that this was a strong contributing factor to the growth of the club. Prior to this players would have to hope to see cross tables presented in Chess Horizons, the USCFs state affiliate for Massachusetts, which up to this time was the only way you could see the results of an event, albeit, usually several months after it had occurred. The club used to have a monthly newsletter to fulfil this role but it had long been discontinued as it was very labor intensive and volunteers always in short supply. Over the next year web site metrics showed the we had a regular spike in visits to our results **every Wednesday morning** and that the number of visits from out of state locations continued to grow. People who didn't even play at the club were apparently interested in seeing our Tuesday night results.
- In 1998 and 1999 attendance jumped from the previous year's average of 35 to 55. In July of 1998 we introduced a third section. The rules of thumb we followed were that we wanted to “protect” the lower players from higher players as well as make it less likely that higher rated players will encounter lower rated players. By offering more sections we were able to reduce the often discouraging experience that occurs in smaller clubs and even in our own clubs past of having once section with a wide range of ratings.
- In 2000 and 2001 our average went to 60 players
- In 2002 our average jumped almost by ten to 69 players
- In 2003 and 2004 we averaged 85 players and our average has been in the 80 to 90 player range ever since. In 2004 we started introducing a fourth section and by 2005 we made it a standard of our program.

An interesting side note about the previous few bullets ... those years we were experiencing our most dramatic growth were the same years where we could find lots of articles about how the internet was spelling the end of over-the-board chess clubs. We knew different. We averaged at least one new player starting at the club each month and during those years almost all had been playing online and wanted to play against “real people” and not on the net or against their computer.

Once we settled on four sections we made a practice of moving the break points each month to appeal to the wide range of ratings of our players. This would allow people to sometimes be at the bottom of a section and other months to be at the top of their section. We also selected the break points based on the trend of ratings we were seeing attend. Sometimes we had a lot of lower rated players for the last six months so we moved the bottom section break points around more. Among other factors was the desire for adults with low ratings to not necessarily be in the same section with new to chess kids. There were many different kinds of minor dis-satisfactions with the section breaks and clearly no way to please everyone. By varying the section breaks most everyone seemed to be able to moderate their dis-satisfaction.

By 2009 varying the section breaks didn't seem very effective anymore. We decided to fix the section breaks to what appeared to be the best break points which would weather six month trend changes. This decision aligned with several of the governing principles of the club namely; to provide consistency in everything we do and to make choices simpler for the players.

V. Keeping a current eye on the numbers

Each month as the event's registration closes for the evening I do some quick data crunching in my head, usually starting by looking at the stack of yellow registration cards piled on the TDs desk. A little later in the evening I then study the pairing charts while waiting for my opponent to make their move. Typically I mentally note most of the following:

- Have we hit our ideal and often typical number of 20 players in each section
- Which sections are above or below 20 players and by how much
- How many players joined this round
- How many players have taken bye's this round
- How many players have arrived late and are they the typical late arrivals

VI. Some new views of our data

After the close of registration for the first round of our July tournament I made a mental note that we have in the Open section:

- 9 Masters
- 5 Experts
- 3 Class players

We have had a number of tournaments in the last year where we've had a strong Open section and my recollection has been that we've had eight or nine masters before. Perhaps it was eight or nine masters and experts, I couldn't remember. This led me to realize that for all the various crunching of numbers we've done we don't have a convenient way of answering questions such as:

- What are the most masters we've had play in an event?
- What would be a reasonable definition of "strong Open section" ?
- What is the strongest Open section we've had?

These questions spurred me on to take revisit the reports we currently generate and consider some new ways of "slicing and dicing it". I have created a collection of new charts which I will summarize below then provide some thoughts about what we can possibly learn from each.

- Modern Era
 1. Section Averages
 2. Balance Between Sections
 3. Balance Between Sections as a percentage
- Modern 80 and standard breaks era
 1. Open Section – All ratings displayed
 2. Open Section – three levels displayed - Masters, Experts and Class
 3. U2000 Section – All ratings displayed
 4. U1700 Section – All ratings displayed
 5. U1400 Section – All ratings displayed
- Open Section participation – All years
 1. All ratings displayed
 2. Displaying only three levels of players – Masters, Experts and Class
 3. Displaying only two levels of players – Masters, Experts
 4. Displaying only one level of players – Masters
- Modern 80 era
 1. Open Section – All ratings displayed
 2. Open Section – three levels displayed - Masters, Experts and Class

VII. The new reports

Each of the graphs displayed in this document can be seen at their full size contained in individual PDFs on the club website.

On the home page there is a link for Demographics on the left edge. This will take you to the complete collection of published data that we maintain.

As of this writing the following link will take you directly to current area, the “Turnout” section of the Demographics on the web site.

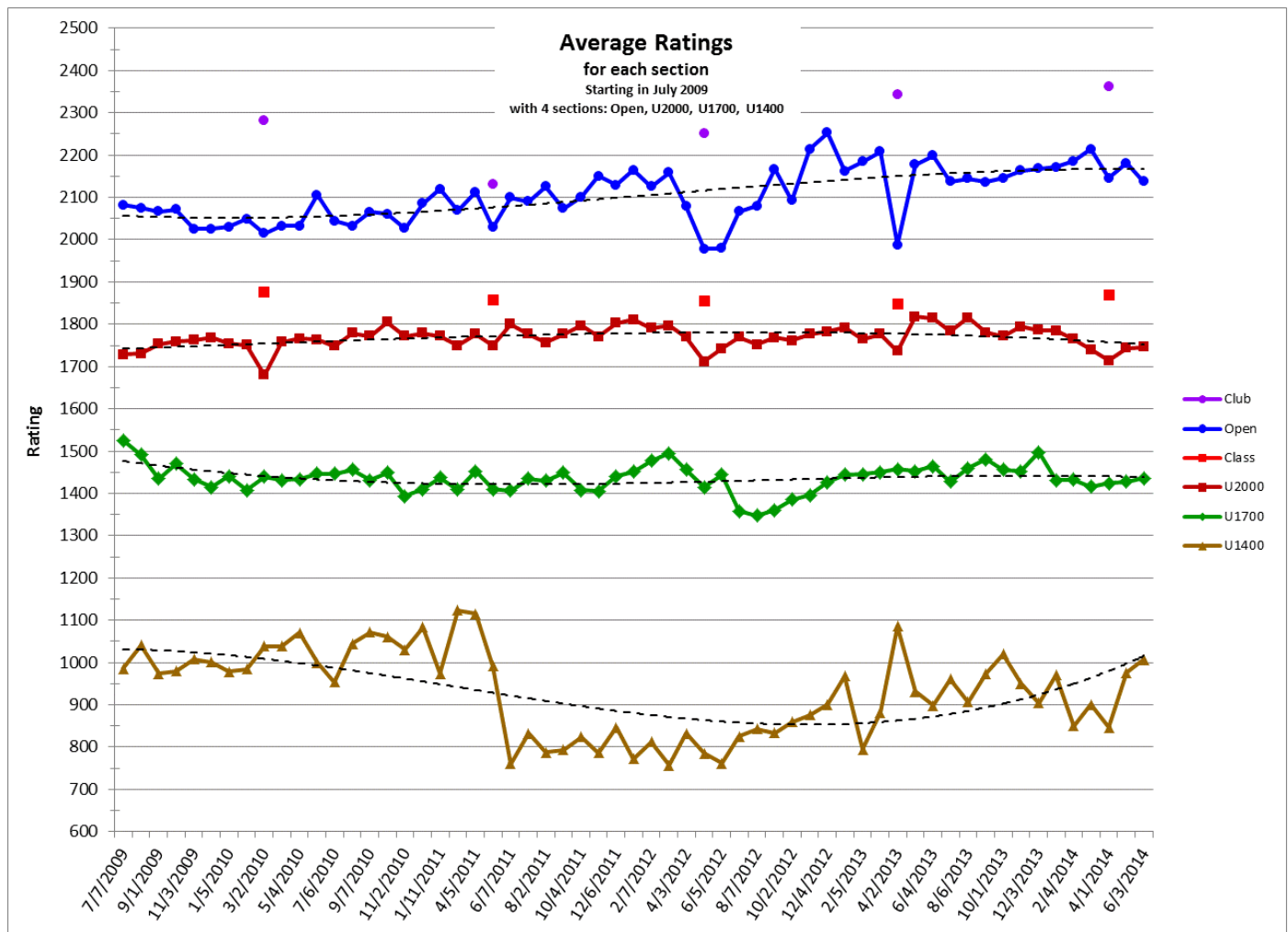
http://www.metrowestchess.org/Community/People/Demographics/Top_Demographics_Turnout.htm

When viewing the charts please be alert to the following:

- There will be some data points that appear to violate the section limit. The ratings used were the pre-event ratings available at the time. Ratings used to have delays of up to almost two months and rating reports used to take a week or two to be processed. By the time we received the rating report back the person’s true rating at the start of the event may indeed have exceeded the section limit. This was and still is the nature of ratings which are snap-shotted several weeks in advance of the start of a month. While rating delays are much less these days there is another cause for such anomalies, namely that events get re-rated at least once a month. This means that the final pre-rating when an event comes back as being rated isn’t so final.
- Some of the charts may be very difficult to read due to the number of dates charted on it. We published these despite their minimal usefulness so that we would know that they are indeed minimally useful and so that we wouldn’t wonder why they seemed to be missing. Modern Era Reports
- Side games are not included in the data
- Re-entries are treated as though they were two different people and thus one person could be counted twice.

A. Section Averages

This chart shows information that we have never seen assembled in such manner and thus provides a new view into our data.



The Open section graph shows a trend line hovering between 2050 and 2150. As the number of Masters who play at the club can and does vary a great deal, there is probably not much to learn from this graph other than the relative strength of the section over time. As with other sections there are typically a number of players who play up into the Open section.

The U2000 graph shows a trend line fitting between 1700 and 1800. This U2000 section has a width of 300 points from 1700 to 1999. An even distribution of ratings would put a third in each the 1700 – 1799, 1800 – 1899 and 1900 – 1999 range and create an average of about 1850. If you jump ahead to the chart showing the U2000 distribution of players you will see from eyeballing the data that U1700 players typically make up about 25% of the players and combined with the 1700 players make up a total of about 50% of the entire section. Note also that there are typically a much larger number of 1800 players than 1900 players. This I believe would account for the low skewing average of about 1750 for the section. These observations support the casual observations that have been made about players wanting to escape playing the players from the lower half of the section below them. This same observation has been made about the U1400 players who play up into the U1700.

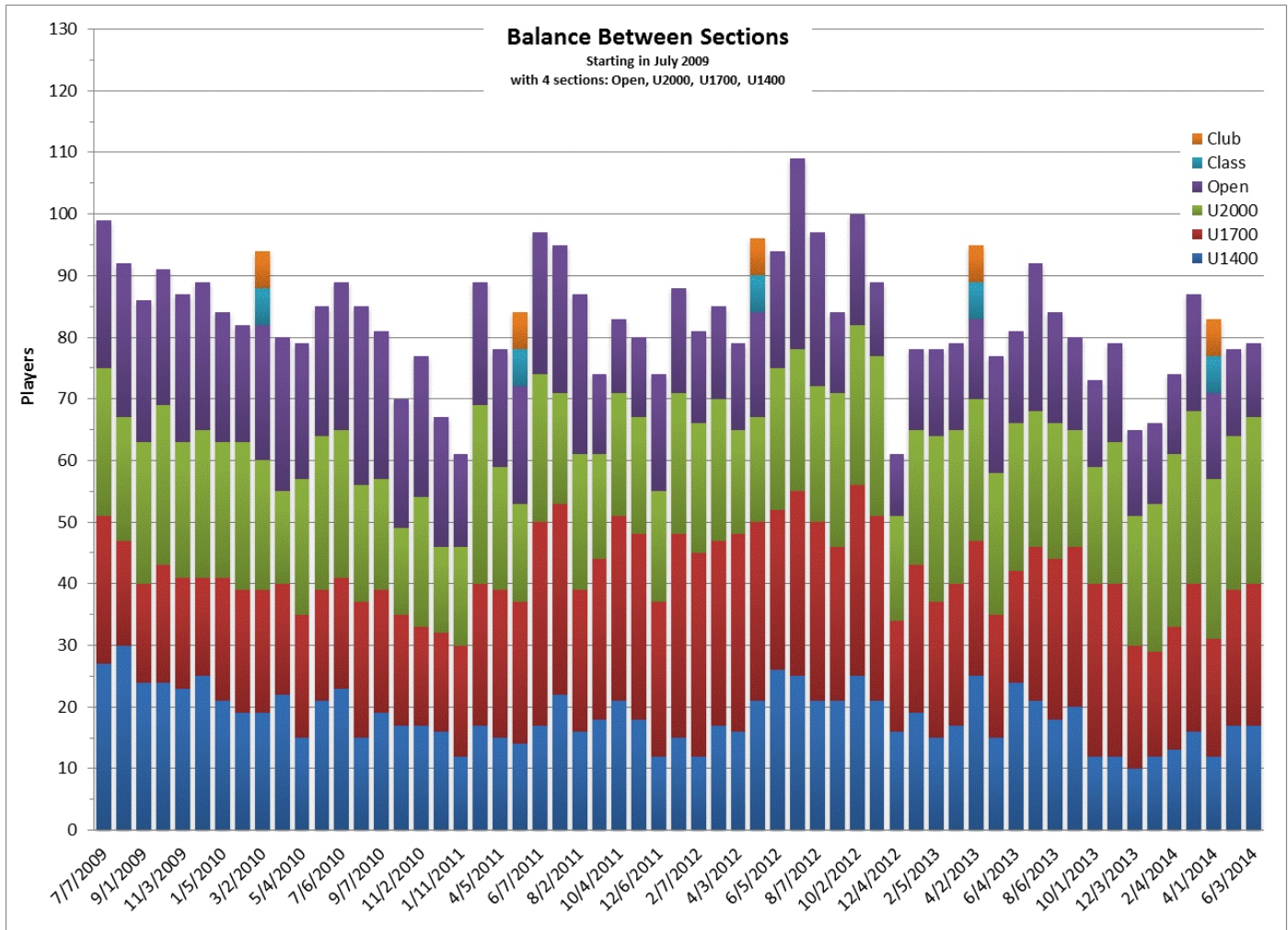
The U1700 shows a trend very close to 1450. This section demonstrates the same traits as the U2000 but with even more lower rated players making up a greater proportion of the section. This I believe accounts for the trend being very close to the floor of the U1700 section.

The U1400 section like the Open section is an opened ended section but where the Open section is open ended in the higher rating direction the U1400 section is open ended toward the lower

rating direction. I believe there is little to learn from the trends in this section other than it reflecting the pool of available players who are currently participating. While the Open section has a natural scarcity of players achieving high ratings, new players in the process of establishing a rating is a potentially limitless pool of players and their number changes with the times.

B. Balance Between sections

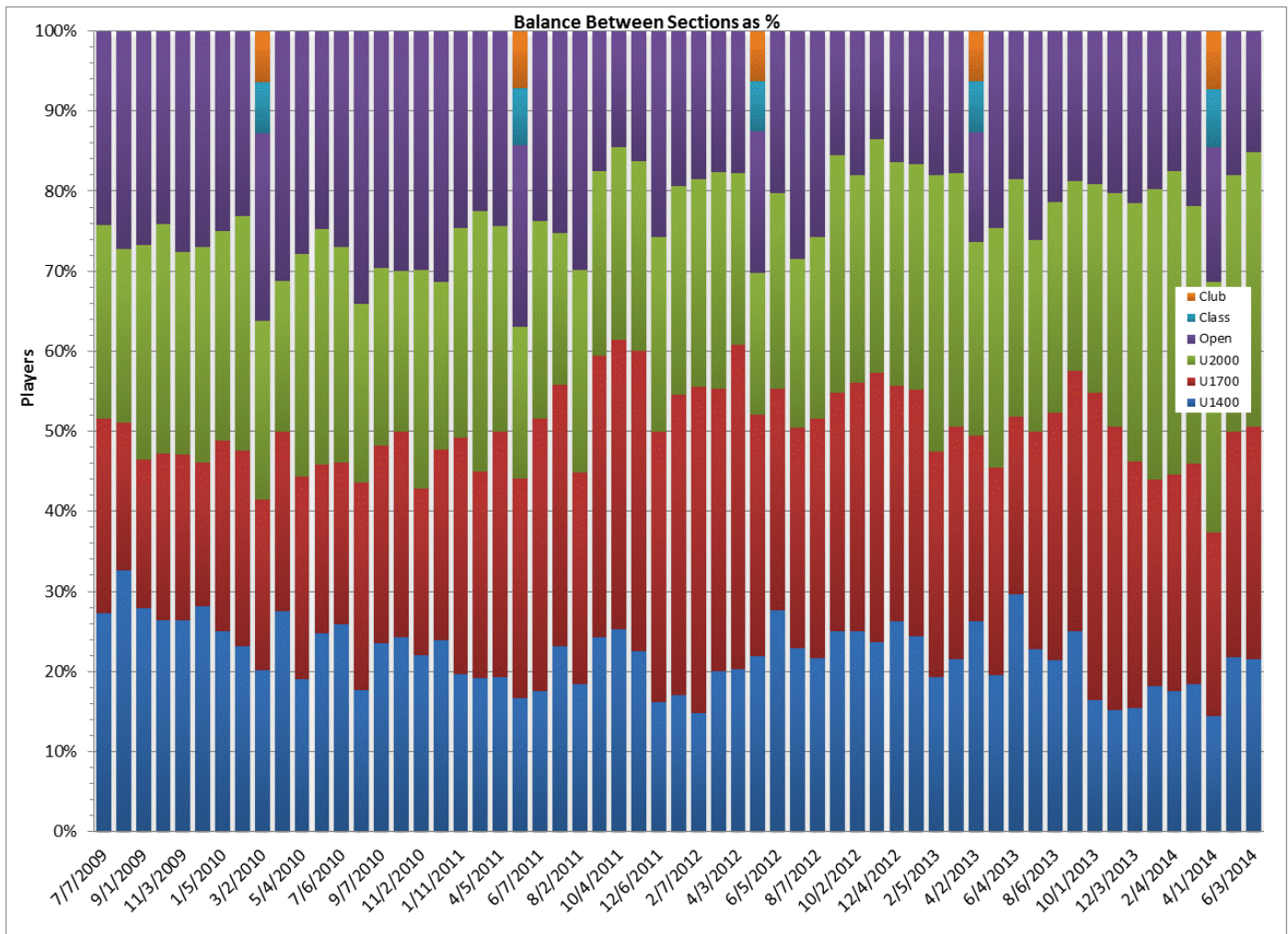
The next two charts show the balance of players between the four sections.



This chart demonstrates the variation that occurs month to month between the sections. There is a casual observation that when the club championships are held more people play up than usual.

The middle sections tend to average closer to 20 players or more while the top and bottom sections tend to have smaller than the 20 players we would hope to see in an equally balanced event. The middle two sections tend to make up for the deficit of the top and bottom sections keeping our average for the event at 80. This chart shows the absolute numbers. The next chart shows the same data but as a percentage of the whole attendance for the event.

C. Balance Between sections as a percentage

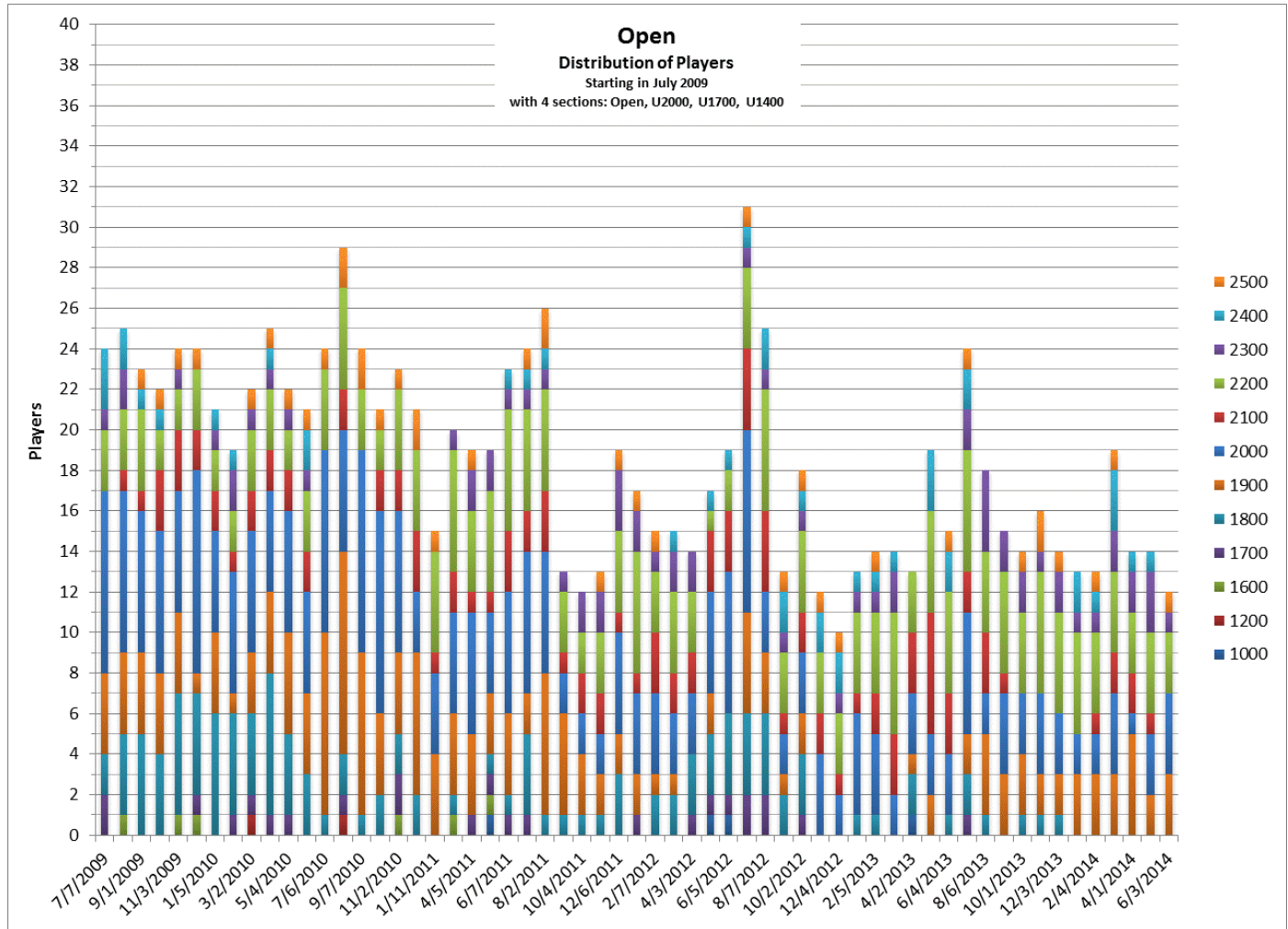


The chart shows the same data as the previous but by percentage of the whole attendance for the event instead of the absolute number of players.

VIII. Modern 80 and standard breaks

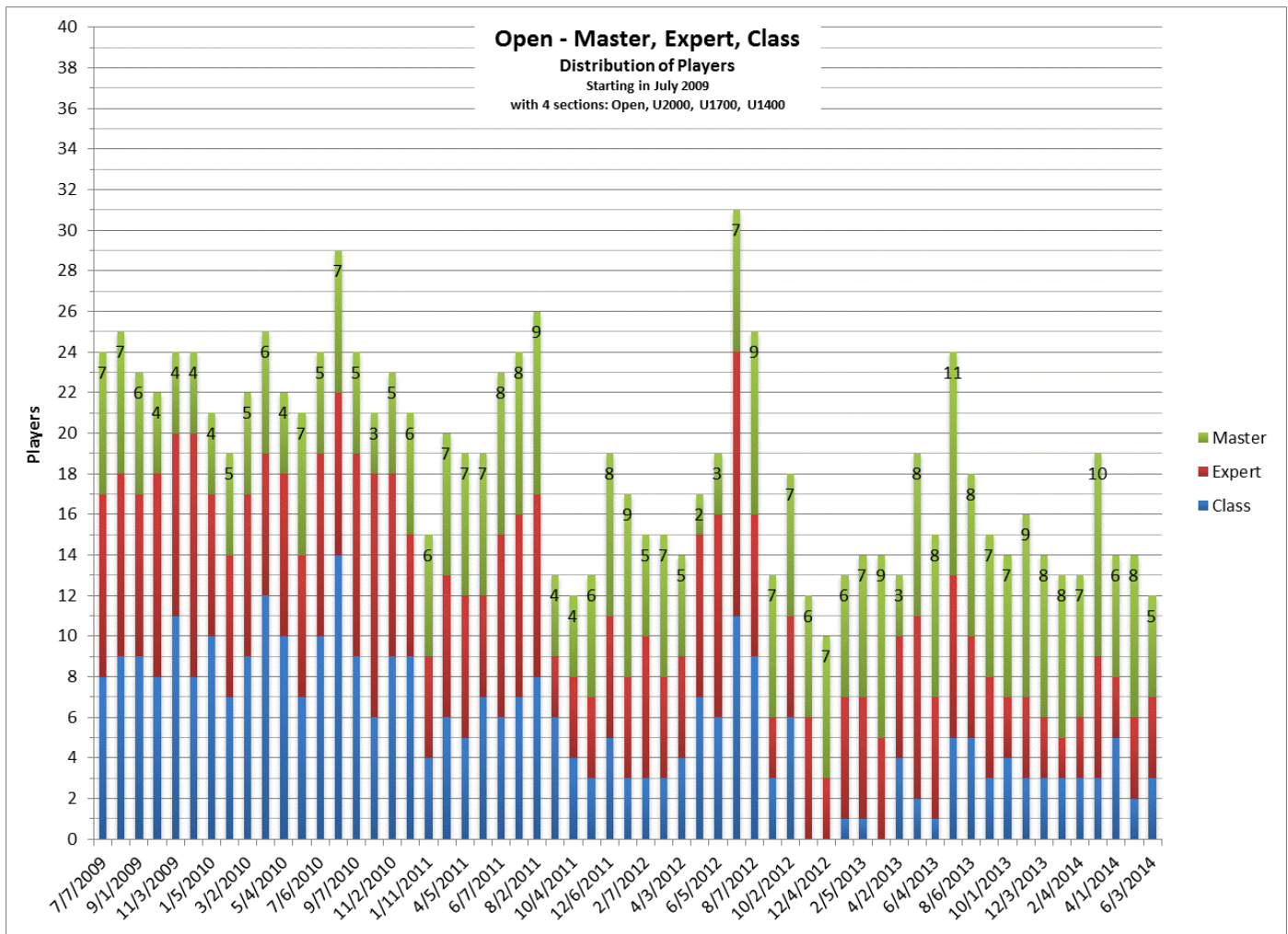
A. Open Section

The next two graphs illuminate the same points as the two just previously discussed. These graphs and the ones to follow in the next section show the same “apples to apples” range of dates, namely from the start of when we fixed our four section breaks at Open, U2000, u1700 and U1400. During this time as the era name indicates we have been running at an average of 80 players each month. This era provides the most consistent view and conditions for us to examine and the most current data as well.



This chart is a smaller time scale and provides a more readable display of modern information. No additional analysis is provided. Refer to earlier sections analysis.

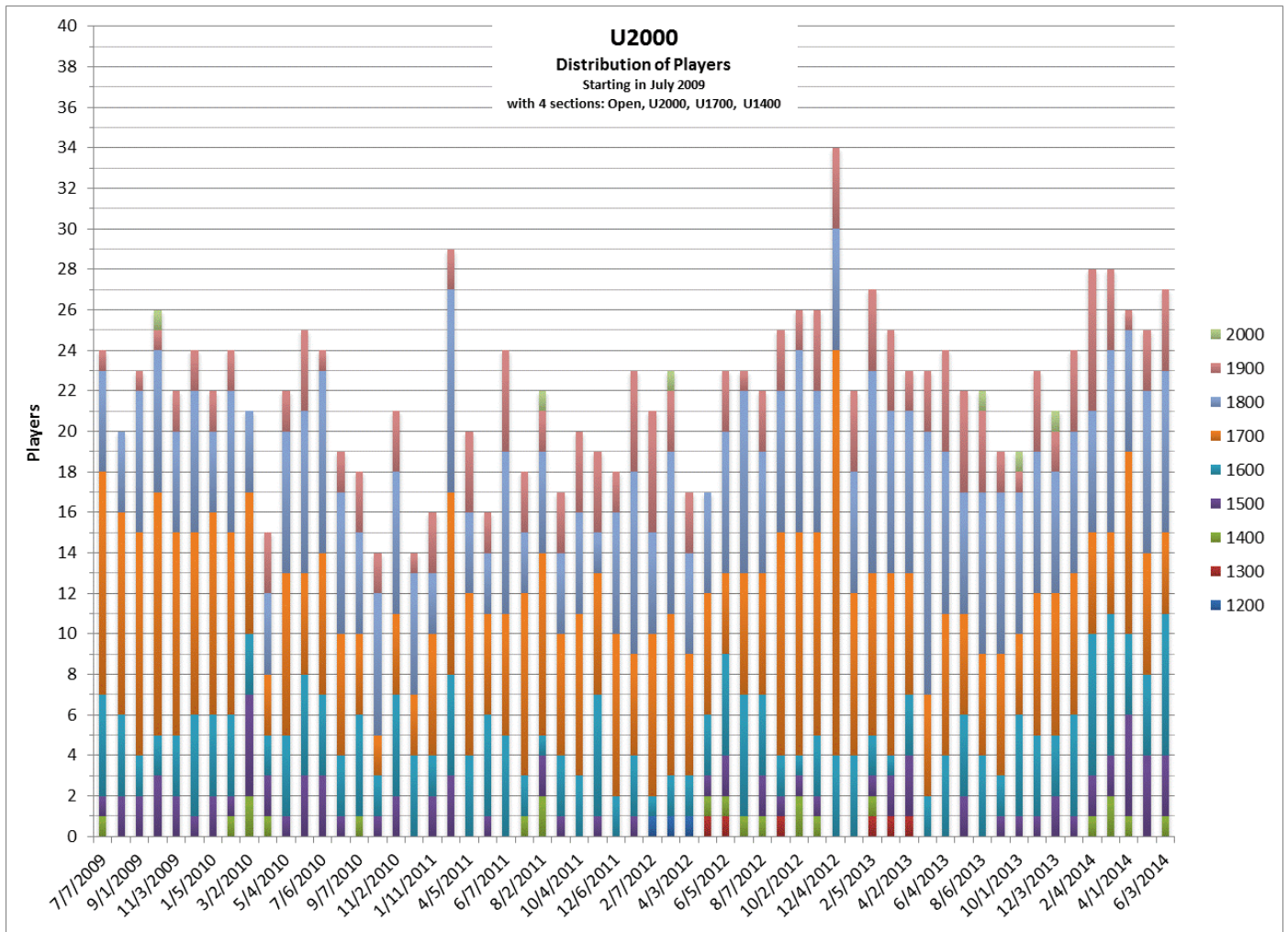
B. Open Section – 3 Levels



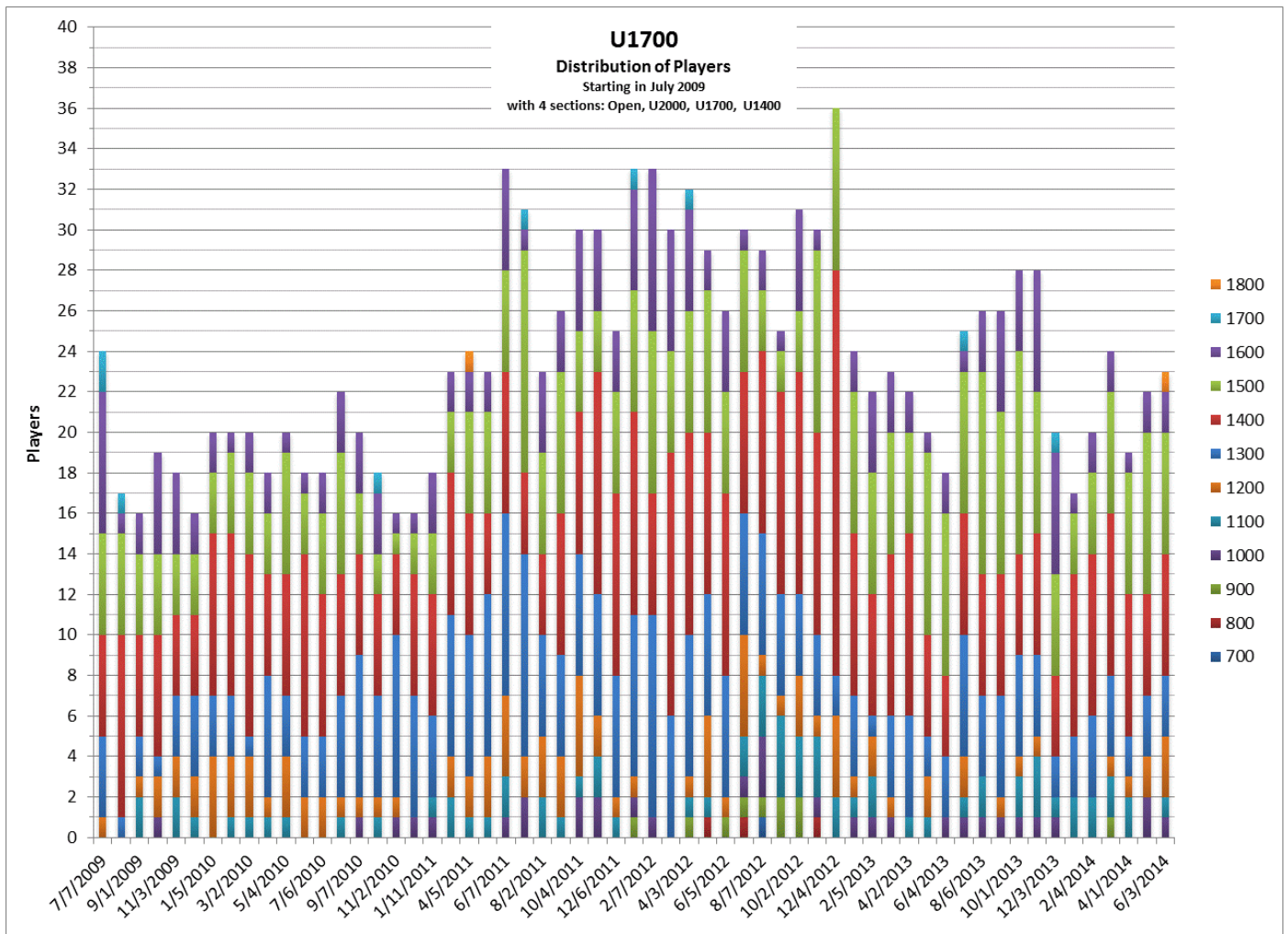
This chart is a smaller time scale and provides a more readable display of modern information. No additional analysis is provided. Refer to earlier sections analysis.

C. U2000 Section

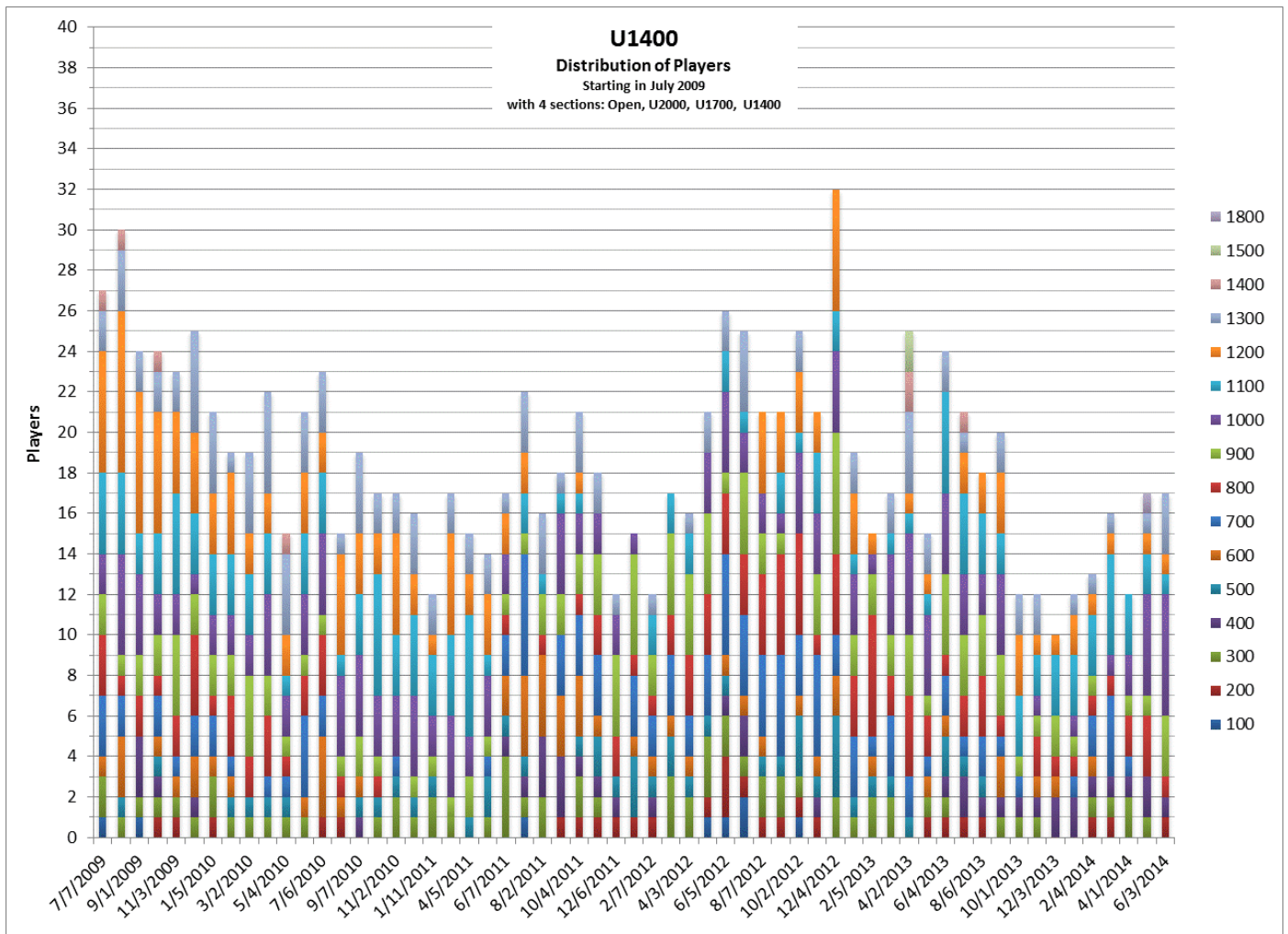
The next three charts show the detailed composition of the non-Open sections. These charts illuminate the change in composition of the sections. No additional analysis is provided.



D. U1700 Section



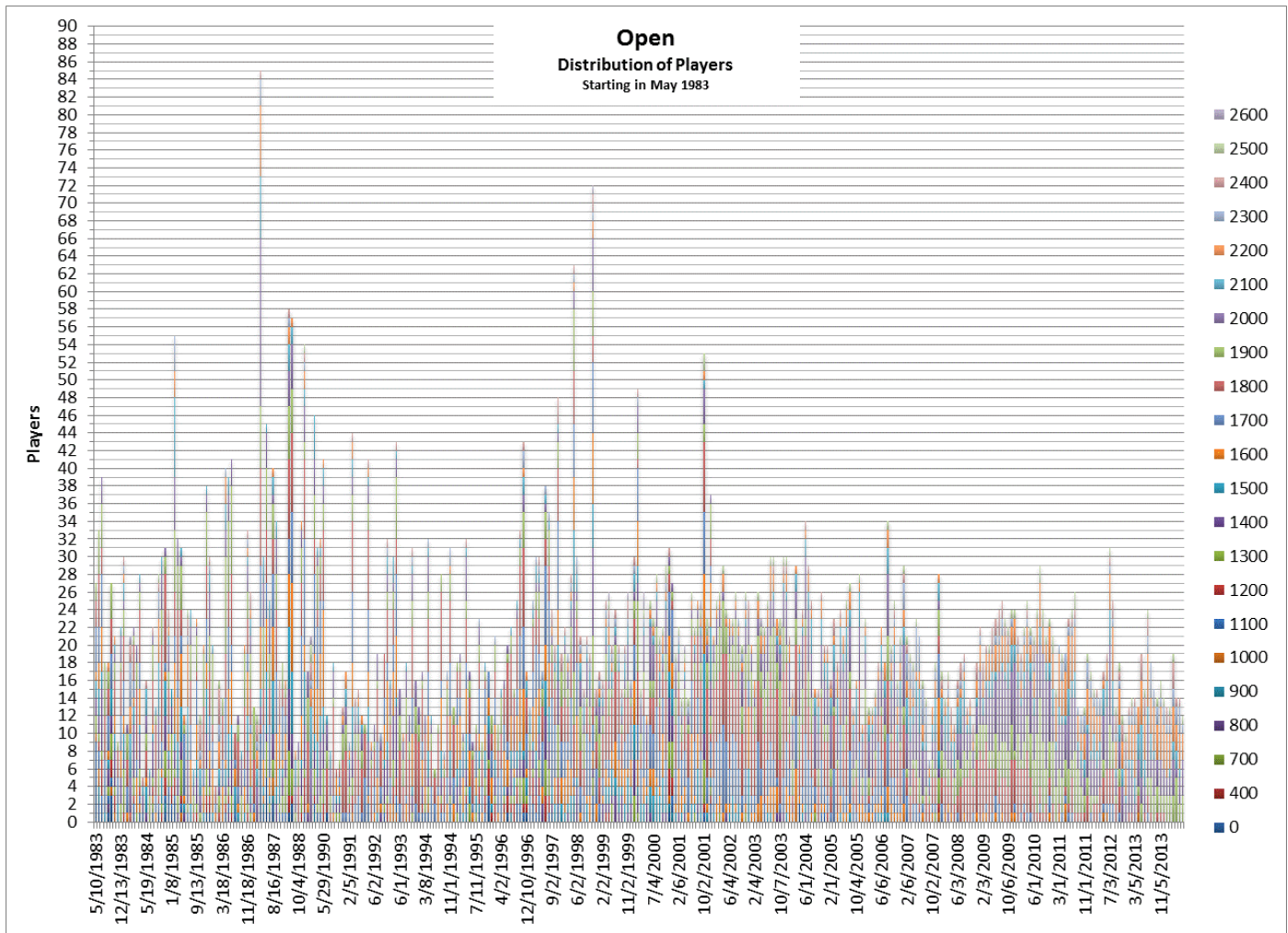
E. U1400 Section



IX. All Years

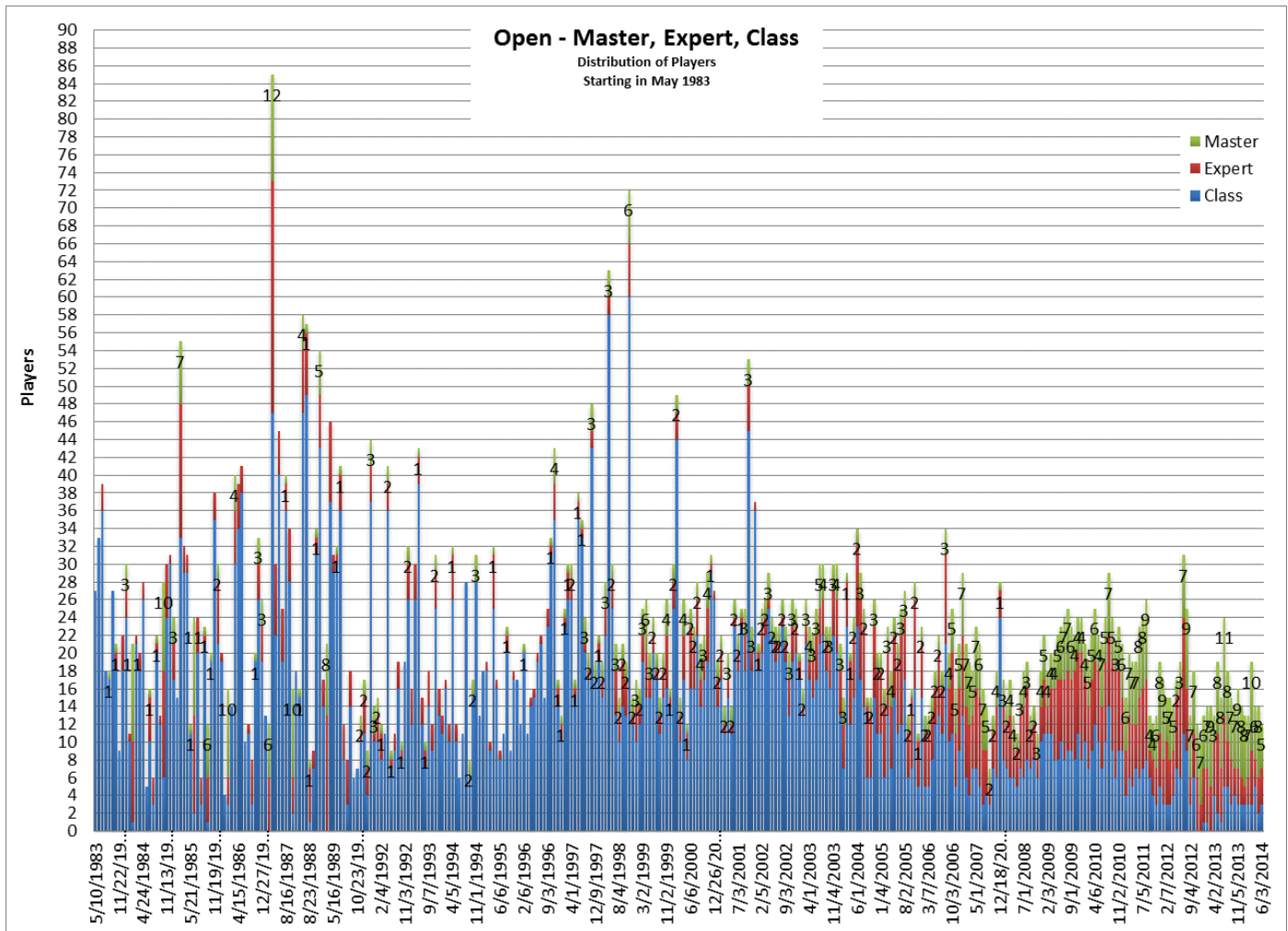
A. Open Section participation

This Chart is simply too difficult to read for details. It does however provide an overview of where peaks of attendance have occurred for the Open section. The early years of the club featured a high proportion of single section events and thus the number of players in the Open represent the entire number of players in the event.



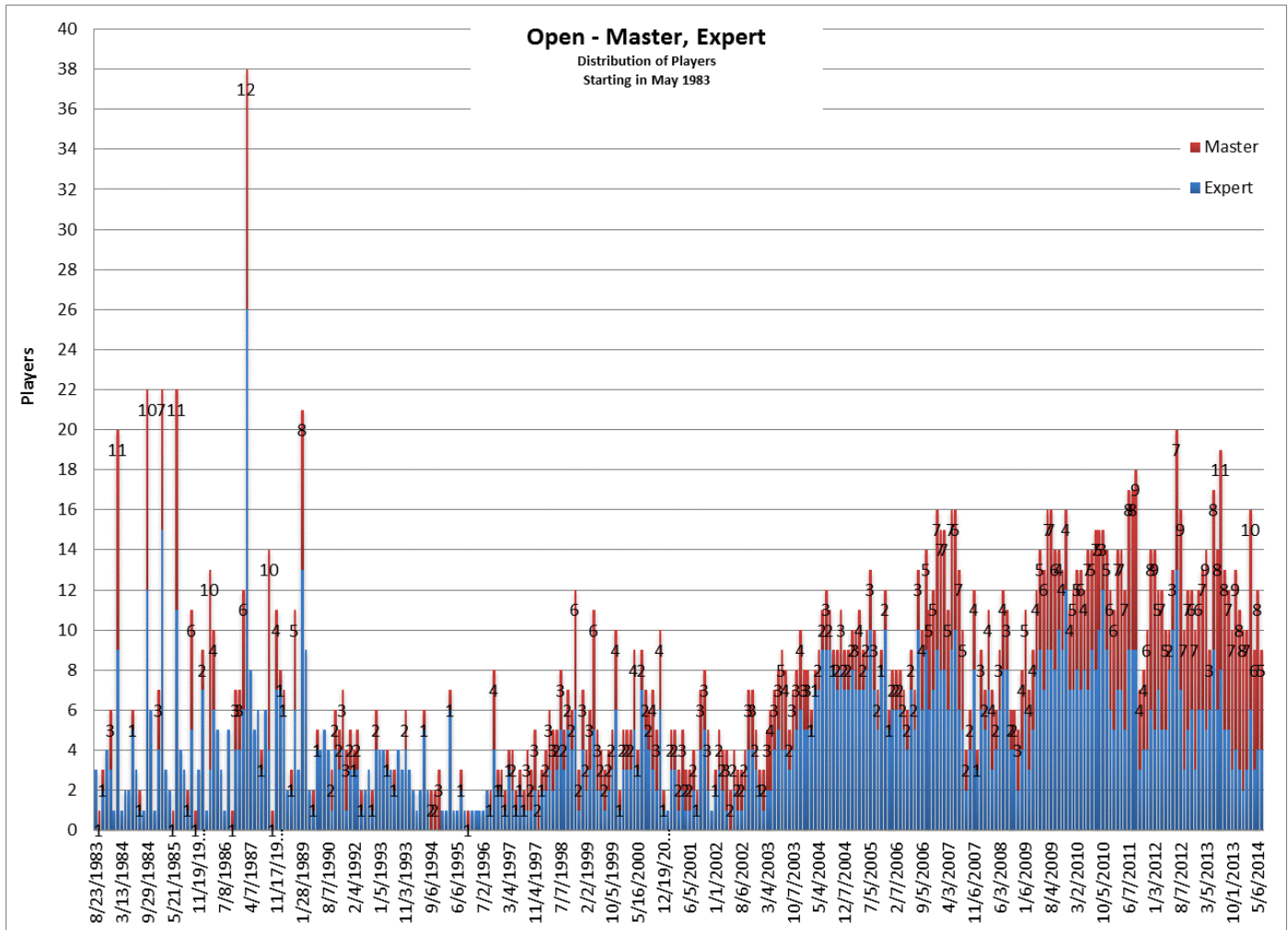
B. Open Section - 3 Levels

Three charts will be discussed in this section. The previous chart showed the breakdown of all ratings in the Open section. The three charts that follow group players into just three groupings, Master, Expert and Class. Each chart displays a diminishing number of these groupings to give better insight into the overall strength and attendance of strong players in the Open section.



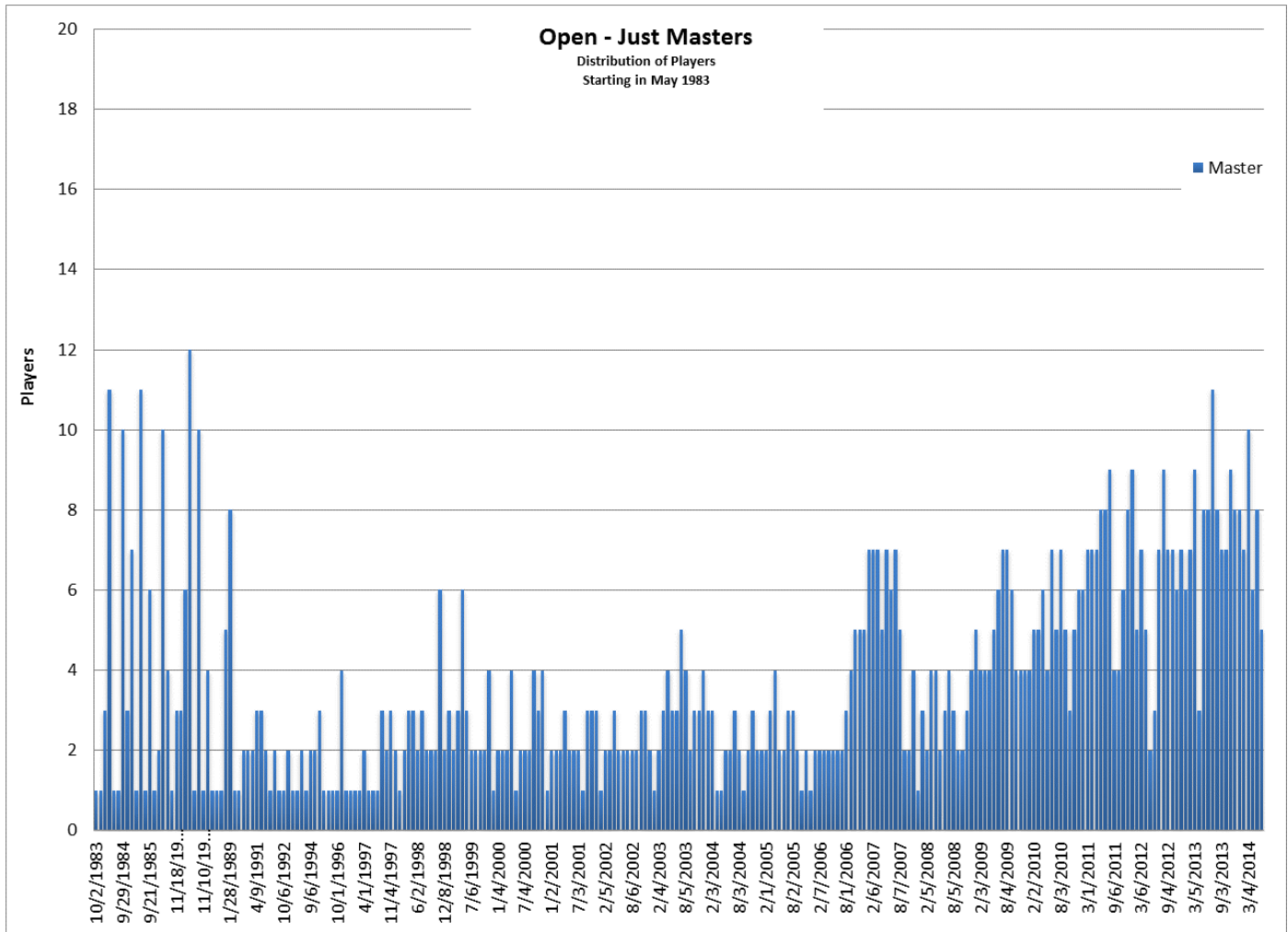
In the chart above you can easily see that class players made up a huge part of the Open section players in the early to middle years but constitute an ever decreasing proportion in the few years. The club went to four sections in 2003 and soon after the number of Experts attending increased as well as Masters. My opinion is that over time players who regularly attended the club began to prefer playing in their own section rather than play up while the number of stronger of players increased. There is always some number of players playing up into the next section but the number seems to be 5 or less players on average going back through 2012.

C. Open Section – 2 Levels



In the chart above the class players have been removed. This graph provides a clearer view of the relative strength Open section through the club's history. Back in the 80s and 90's there were simply a lot fewer masters around. As noted earlier, after the first round of this month's even we have 9 Masters and 5 Experts playing for a total of 14 players. 14 players is certainly at the higher end of the range on the graph above but not exceptional

D. Open Section – 1 Levels

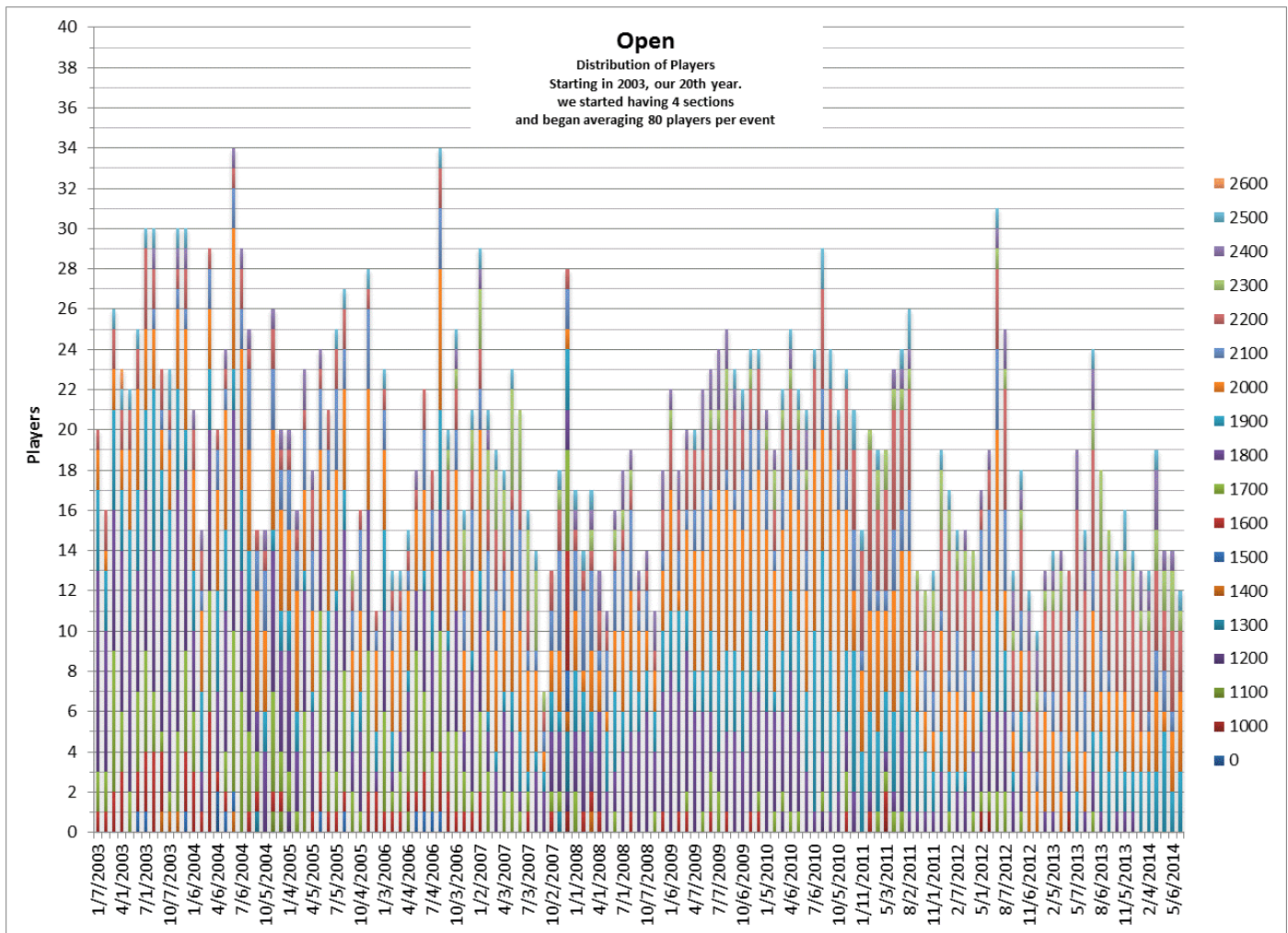


The chart above helps answer the question about what the record number of Masters are as only Masters are displayed. As noted earlier, after the first round of this month's even we have 9 Masters playing. By my count there are only nine events with 10 or more Masters so this month's event may prove to be one of the highest number of Masters event.

X. Modern 80 era

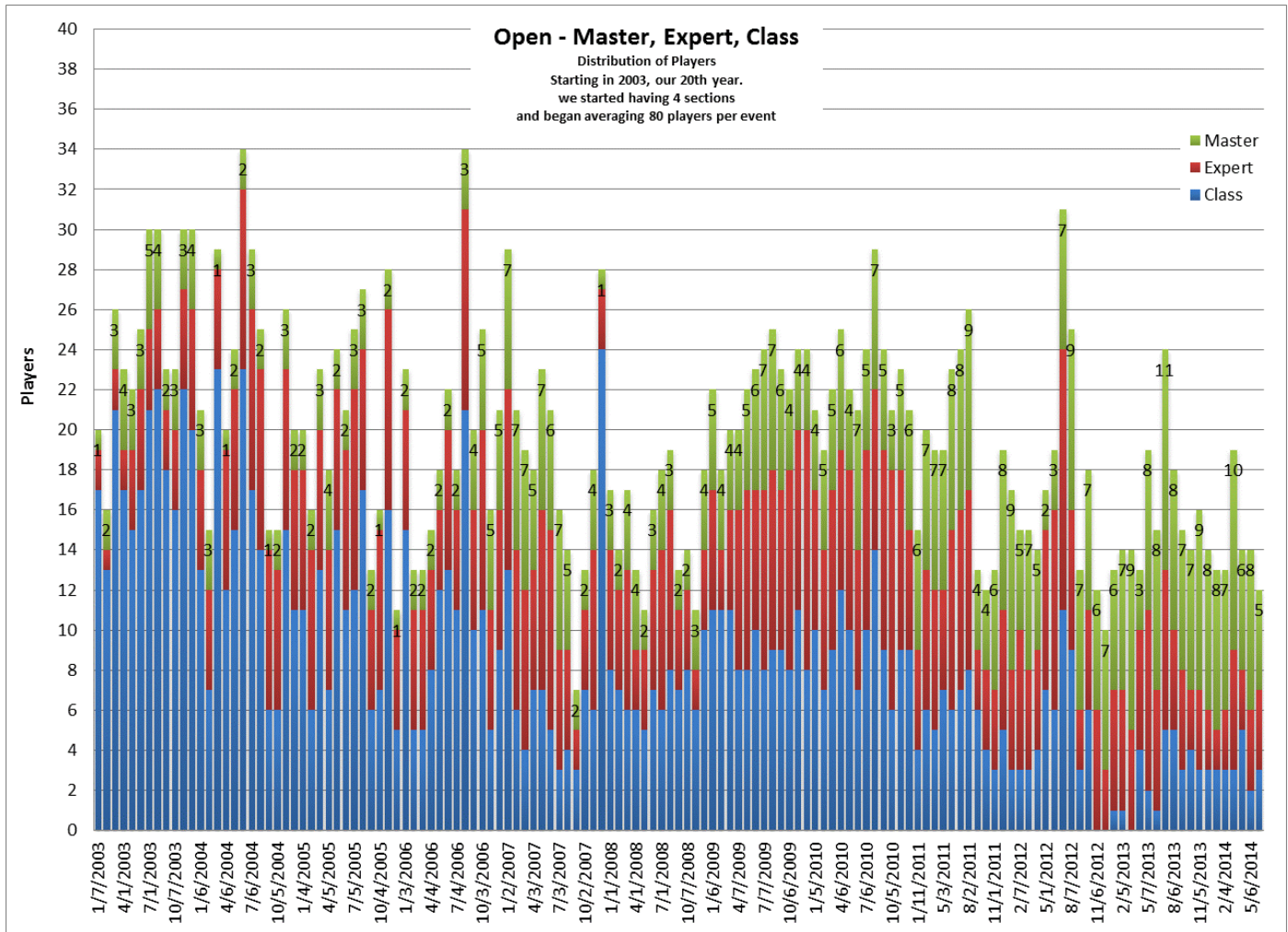
A. Open section

Two graphs will be discussed in this section. Both of these graphs cover only the Modern 80 Era which is defined as starting in 2003. This eliminates the first 19 years of data and thus allows us to see much greater detail than most of the earlier charts discussed.



The chart above shows the composition of the Open section since 2003. From 2003 until July of 2009 our regular time control events had three sections whose break points moved from month to month. It is still difficult to see an indicator or relative strength or proportions of class players so we have the next chart to aid us.

B. Open section – 3 Levels



The chart above doesn't really provide any insight we haven't already discussed, it simply shows the data more clearly than previously. Here it is easier to observe the 5 or less players playing up since 2012 that was discussed earlier. Here it is easier to see that six or less players typically played up going back to the start of 2011. Yes there a few exceptions, thus the use of the word typical.

XI. Answering the Questions

Earlier I stated that this current look at our data was prompted by some questions that came to mind after the first round of this month's event where we have 9 Masters, 5 Experts and 3 Class players. The questions are now discussed below.

A. What are the most masters we've had play in an event?

Using the "Open Section – 1 Levels" chart to see just the Master across all years of the club, we see the following:

- 12 Masters – 1987-03
- 11 Masters – 1982-02, 1985-06, 2013-07
- 10 Masters – 1984-09, 1986-02, 1987-09, 2014-03,
- 09 Masters – 2011-08, 2012-01, 2012-08, 2013-03, 2013-11

As noted earlier, after Round 1 of this month's event we have 9 masters.

The event that saw 12 Master, 1987-03, was a special event sponsored by the club, The NE Amateur team Championships. This event drew participants from all of New England.

Based on current participants and the high level of competition in the Open section we will continue to see nine or ten Master events and will likely tie or break the record of 12 Masters soon.

B. What would be a reasonable definition of "strong Open section"?

When I created these new views of our data I had been thinking that several approaches would be considering and created the charts to examine the data for each.

1. Average of ratings in the Open section
This might be the simplest approach as it accounts for dilution of the strong players by weaker players.
2. Ratio of Masters, Experts and Class players in the Open section
Using the chart Open 3 Level we can see the composition of the open section. I'm not confident that this approach works well without taking into consideration the number of participants. For example, would an event with 5 Masters, 5 Experts be a stronger section than an event with 3 masters and 9 experts? Does this lead us to consider using averages?
3. Count of Masters, Experts and Class players in the Open section
Again I'm not sure how we could apply this approach fairly. I think it runs into about the same issues as the Ratio approach.

C. What is the strongest Open section we've had?

As we are probably aware, the same data can be interpreted different ways. Perhaps someone can propose a better or more refined approach than I was able to think of. I'm going to choose the average of ratings approach as its simple and accounts for dilution of lower rated players.

From the Modern 80 with fixed section era: the strongest event was the 2012-12 event. That event had 7 Masters and 3 Experts.

I did not create a chart for the average across all events but I present the data below. I've selected only the events where the average was 2200 or above.

Average	Section	Date
2361	Club	04/01/14
2342	Club	04/02/13
2282	Club	03/02/10
2253	Open	12/04/12
2251	Club	05/01/12
2245	Open	09/21/85
2238	Open	09/19/87
2232	Open	01/28/89
2231	Open	02/01/86
2229	Open	02/04/84
2225	Open	01/31/87
2214	Open	11/06/12
2214	Open	03/04/14
2208	Open	03/05/13

From the table above we see that our strongest sections were not Open sections but our Club Championships. Our Club Championships are by invitation only. To qualify you must be a top earner Championship points during the preceding year. The points are earned by placing first, second or third in the Open, U2000 or U1700 sections.

The 2012-04 event had 6 Masters, 3 Experts and 5 Class players for a total of 14 players.

Another observation from the above data is that our strongest sections have occurred in the last five years. The next clustering of strong events occurred in the mid to late eighties. The remaining other events with an average of over 2200 also occurred in the last five years. Of the fourteen events more than half of them were in the last five years.

D. Summary of the answers

- Most Masters:
 - Most - 12 Masters – **1987-03**
 - In the Modern era - 11 Masters - **2013-07**
- Definition I selected of strongest section – Highest Average of players ratings
- Strongest:
 - Club Championships - **2014-04** avg. rating **2361**
 - Strongest Open section – **2012-04** avg. rating **2253**