**Predicting The Ideal Time To Purchase Allergy Medication**

For The Albuquerque Area

Matthew Letter

University of New Mexico: Department Of Computer Science

Big Data

Albuquerque, New Mexico

Mletter1@unm.edu

*Abstract*

Medication waste is a growing problem in the United States. This paper addresses ways to stem allergy medication waste, for the Albuquerque area, through predicting what months have high pollen counts and what factors contribute to pollen levels. The author of this research believes outlining stimuli related to pollen levels can help with medication purchase determination. Data coinciding with pollen level and correlating factors are used to make this determination. If we follow the methods outlined in this paper for choosing when to purchase allergy medication we may very well avoid a large amount of medication waste and improve our ability to predict bad pollen years for the Albuquerque are.

The data yields that the highest contributing factor to pollen levels in the Albuquerque area is rainfall precipitation. Drought years (years with low rainfall) show a drastic drop in pollen level implying that pollen levels are directly related to rainfall totals. In this was a prediction that a drought year yields low pollen counts has been made using tableau. Furthermore, on average, the start of the allergy season in the Albuquerque area is February with the elevation of pollen levels such as juniper. The peak of allergy season is April with Mulberry being the strong overarching pollen producer over all area plants tested. By November the pollen levels have bottomed out therefor it is recommended that this is the best month for purchasing allergy medication, as allergy medication expire in 12 month increments (such as 1 year, 2 year…). Furthermore, for years with droughts it is recommended that you wait to buy your medications and see if you allergies are sever enough to warrant purchasing medication as the level of pollen in the air is on the order of magnitude 10 times lower than years without draught.

Keywords; Pollen, Albuquerque, Purchasing, Allergy, Medications, calendars, Weather.

# Introduction

People spend billions of dollars a year trying to control their allergies [2]. Unfortunately a large amount of allergy medication are thrown away because people buy their allergy medication at the wrong point of time during the year, which leads to expiration before they can be use. The goal of this study if to give people in the Albuquerque area a way of determining when to buy there allergy medication so as to not let them go to waste. Furthermore, this paper looks and pollen trends in Albuquerque to help people with specific allergies predict when to buy their medications.

Items that are looked at include: how do pollen levels stack up based on the plant type, what months of the year do certain plants pollenate, what weather factors contribute to pollen levels, and what insight can be gleaned from the weather. These factors were all used to come up with a set of conclusions that help us determine the proper time of the year to by medications based off of insights from the collected pollen data. Furthermore these insights give us the ability to predict what year may have high levels of pollen relative to the norm.

# Related Work

## Allergenic Pollen and Pollen Allergy in Europe

Allergies have been impacting people all over the world and weather conditions appear to play a significant roll in the ability of plants produce pollen [1]. Data obtain from aerobiological studies and allergological investigation; have contributed to the creation of pollen calendars with the approximate flowering period of plants in a given region. It has been noted that pollen data correlates with weather conditions for given regions. Allergens such as pollen have contributed to degrading the quality of life of people suffering from the many respiratory conditions, such as asthma.

*2.2 A Revised Nomenclature for Allergy: An EAACI Position Statement from the EAACI Nomenclature Task Force*

The European Academy of Allergy and Clinical Immunology (EAACI), an association promoting basic and clinical research, reports a revised nomenclature for allergic reactions based on the present knowledge of the allergy mechanisms. Allergic reactions can express themselves in many different bodily organs through many different mechanisms. According to [2], an allergy, a specific type of hypersensitivity, is a clinical reaction in which an immunologic mechanism is proven or strongly implicated. An allergy can be antibody or cell mediated. In an effort to avoid misunderstandings between patients, physicians, and their colleagues, a very clear designation and nomenclature of the allergic disorders must be adhered to.

*2.3 A Holocene pollen record of persistent droughts from Pyramid Lake, Nevada, USA*

Pollen and algae preserved in the sediments from Pyramid Lake, Nevada, provide evidence for periods of persistent drought. Pollen levels increase when the ground has sufficient saturation for plant reproduction. The level of pollen for a wet year will be increased from that of a dry year. This fact was used to analyze core samples from Pyramid Lake, Nevada, showing that 7600 to 6300 cal yr B.P., was the driest period of the Holocene age [3].

*2.4 Superiority of an Intranasal Corticosteroid Compared With an Oral Antihistamine in the As-Needed Treatment of Seasonal Allergic Rhinitis*

Daily use of nasal spray corticosteroids or histamine receptor antagonists have proven to be effective in the treatments of allergic rhinitis. One thing that is noted is that the use of histamine receptor antagonists should be taken as needed. Therefor it is important to keep track of when to buy your allergy medications. This study shows that the as-needed intranasal corticosteroid outperformed antihistamines when controlling allergic rhinitis [4].

# Design and Approach

## Obtaining and Cleaning the Data

Data was obtained from the City of Albuquerque public data set website, ABQ data [5]. The data was presented in the form of an xml file with 70,000 data points. This xml file was parsed and clean for insertion into relation SQL database implemented with the MySQL engine. A second set of weather data for the Albuquerque was obtained. This data set had 250,000 data points. This data set was obtained as a CSV file from Wunderground and was parsed, cleaned, and inserted into the MySQL database using java [6]. At this point the data was ready for manipulation, interpretation, and prediction. The main reason to upload these to data sets into a MySQL database was for that ability to run inner joins on the data off of the date attribute.

## Data Manipulation Tools

Four different data tools were used for manipulation, prediction, and presentation of the data. The first tool used was Ubiq, which is an Analytics tool that integrates with MySQL. This tool was chosen because of it powerful filters and functions and its ability to represent joins on a bar chart. The second analytic tool used was tableau. Tableau is an analytics prediction tool. Tableaus prediction tool was used for its ability to handle multi thousand line query results, as Ubiq can only handle up to 1000 rows of data. The third manipulation tool was java. Java was used for parsing and cleansing of the data. The final data manipulation tool was MySQL. MySQL is a key player for analyzing big data because it can be used to leverage operations such as joins and the SQL language itself for key insights into Big Data sets.

## Aproach and Design

First a determination was made on whether to use MongoDB or MySQL. MongoDB has the advantage of map reduce functionality, which is ideal for reducing the size of the data. MySQL was chosen over MongoDB for the ability to run joins on data tables, which puts the data sets in the context of each other. Next was determining a language to write the parsers in.

Three languages were considered Python, Java, and C++. Python has many libraries for parsing, machine learning, and connecting to databases. C++ has many of the same feature capabilities as python but you have to worry about memory management. Finally java was looked at and chosen for the language to implement the parsers. Java was chosen manly because of the SAXPARSER library. The sax parser is designed for parsing XML files, as the data collected from the

Albuquerque data warehouse was, in the for of XML. The sax parser took in the XML pollen data and cleaned the attributes subsequently putting them into an SQL database [5]. The second parser was for parsing the CSV data file obtained from Wunderground that contained all the weather data for the Albuquerque area [6]. This parser took in the CSV file and cleaned the data for proper representation to the form of a relational database. Java was chosen to implement this because it was already being used for the Sax Parser. After the data was in the MySQL database it was ready for analysis Using Ubiq and Tableau.

Once the data was in the MySQL database Ubiq was used for initial discovery and insight. The MySQL database was connected to Ubic through an ssh tunnel. Once Ubic had the connection to the database manipulation of the data was done. The first goal was to obtain a representation of the average pollen counts for each month, and a weather data trend line comparison later on. Once an average data plot was produced other plots on various conditions needed to be produced. Plots on pollen level per tree and year are investigated. Leading to

# Results

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

## Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

## Units

* Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
* Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
* Do not mix complete spellings and abbreviations of units: “Wb/m2” or “webers per square meter”, not “webers/m2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
* Use a zero before decimal points: “0.25”, not “.25”.

## Equations

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in

 

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

## Some Common Mistakes

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum **0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
* In American English, commas, semi-/colons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively”.
* In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
* Do not confuse “imply” and “infer”.
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”.
* The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [7].

# Discussion

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper.

## Authors and Affiliations

The template is designed so that author affiliations are not repeated each time for multiple authors of the same affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization). This template was designed for two affiliations.

### For author/s of only one affiliation (Heading 3): To change the default, adjust the template as follows.

#### Selection (Heading 4): Highlight all author and affiliation lines.

#### Change number of columns: Select Format > Columns >Presets > One Column.

#### Deletion: Delete the author and affiliation lines for the second affiliation.

#### For author/s of more than two affiliations: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select Format > Columns > Presets > One Column.

#### Highlight author and affiliation lines of affiliation 1 and copy this selection.

#### Formatting: Insert one hard return immediately after the last character of the last affiliation line. Then paste the copy of affiliation 1. Repeat as necessary for each additional affiliation.

#### Reassign number of columns: Place your cursor to the right of the last character of the last affiliation line of an even numbered affiliation (e.g., if there are five affiliations, place your cursor at end of fourth affiliation). Drag the cursor up to highlight all of the above author and affiliation lines. Go to Format > Columns and select “2 Columns”. If you have an odd number of affiliations, the final affiliation will be centered on the page; all previous will be in two columns.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1”, “Heading 2”, “Heading 3”, and “Heading 4” are prescribed.

## Figures and Tables

### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

1. Table Type Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

a. Sample of a Table footnote. (Table footnote)

We suggest that you use a text box to insert a graphic (ideally 300 dpi), with all fonts embedded) because, in an MSW document, this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord pull-down menu, select Format > Borders and Shading > Select “None”.

1. Example of a ONE-COLUMN figure caption.

Please see last page of this document for AN EXAMPLE of a 2-COLUMN Figure.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

## Footnotes

Use footnotes sparingly (or not at all) and place them at the bottom of the column on the page on which they are referenced. Use Times 8-point type, single-spaced. To help your readers, avoid using footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence).

# Conclusion

You must submit the IEEE Electronic Copyright Form (ECF) per Step 7 of the CPS author kit’s web page. THIS FORM MUST BE SUBMITTED IN ORDER TO PUBLISH YOUR PAPER.

Please see Step 9 for ordering reprints of your paper. Reprints may be ordered using the form provided as <reprint.doc> or <reprint.pdf>.

##### Acknowledgment

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression, “One of us (R.B.G.) thanks . . .” Instead, try   
“R.B.G. thanks”. Put applicable sponsor acknowledgments here; DO NOT place them on the first page of your paper or as a footnote.

##### References

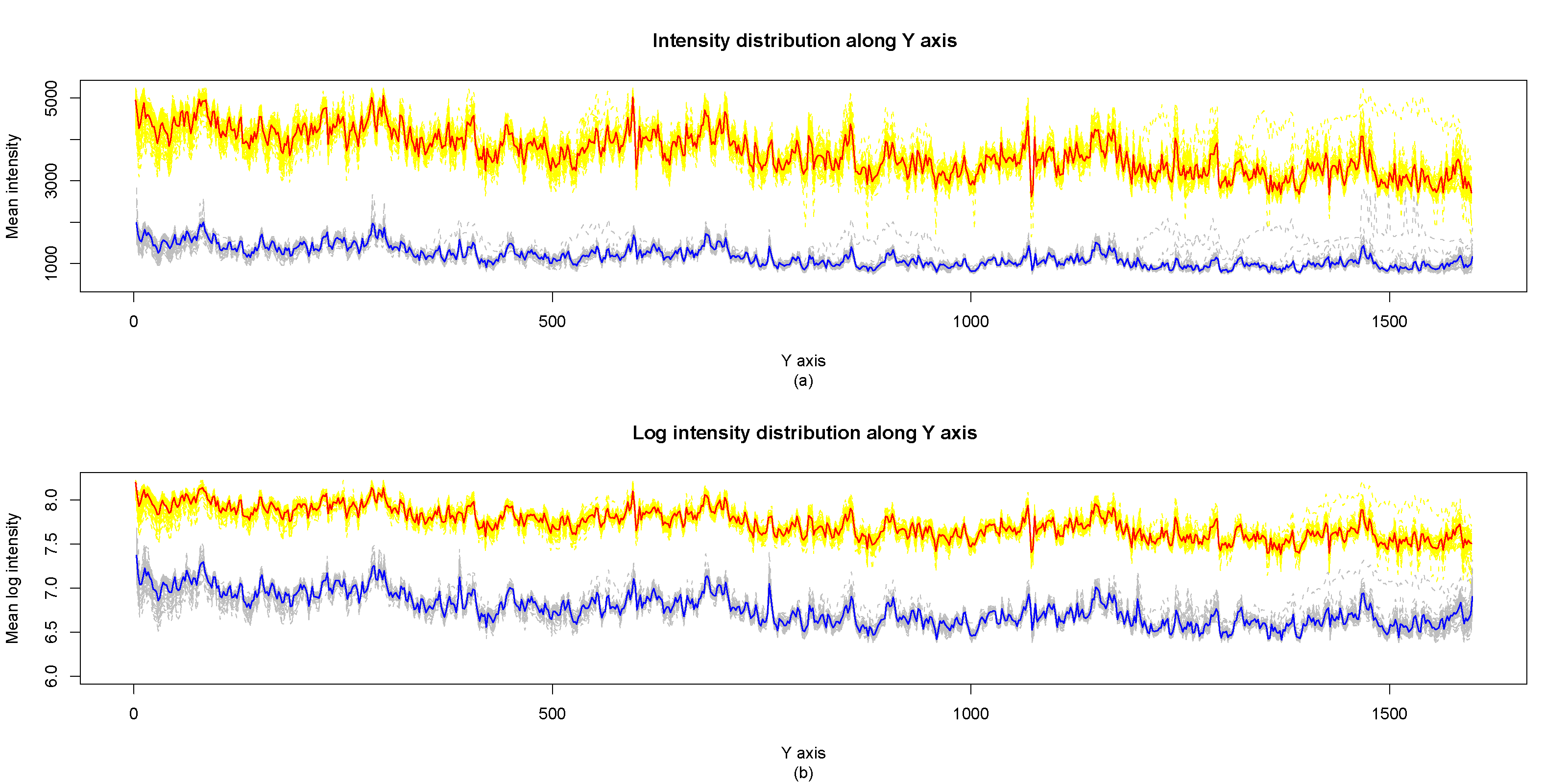
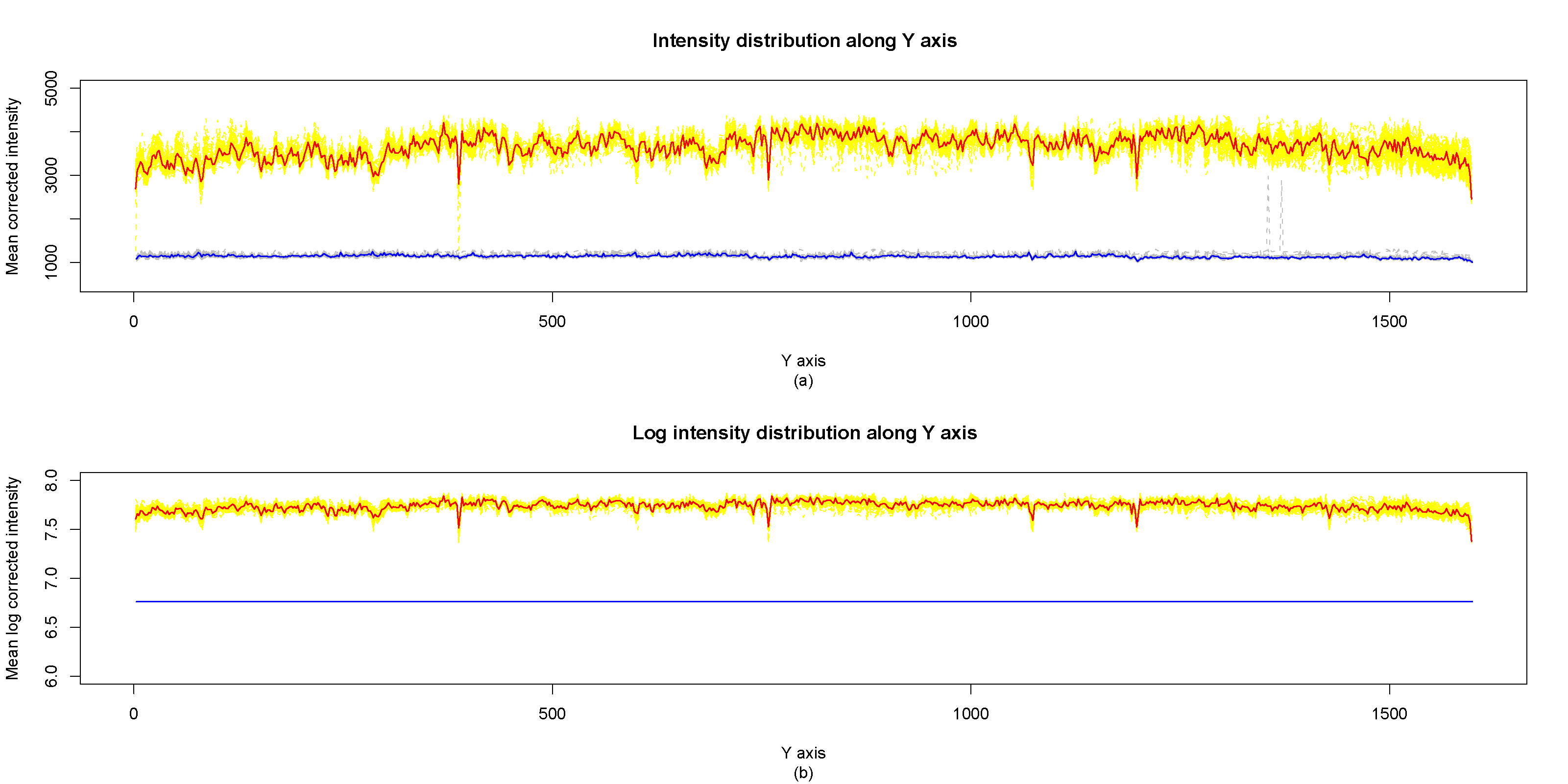
List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [1]. Where appropriate, include the name(s) of editors of referenced books. The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first . . .”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. D’Amato, G., et al. (2007), Allergenic pollen and pollen allergy in Europe. Allergy, 62: 976-990. Doi: 10.1111/j. 1398-9995.2007.01393.x
2. Mensing, A. M., et al. (2001), A Holocene pollen record of persistent droughts from Pyramid Lake, Nevada, USA. Doi:10.1016/j.yqres.2004.04.002
3. I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
4. Kaszuba SM, Baroody FM, deTineo M, et al. Superiority of an intranasal corticosteroid compared with an oral antihistamine in the as-needed treatment of seasonal allergic rhinitis. Arch Intern Med. 2001;161(21):2581–7. [[PubMed](http://www.ncbi.nlm.nih.gov/pubmed/11718589)]
5. "ABQ Data." City of Albuquerque. City of Albuquerque, n.d. Web. 06 Oct. 2014. <http://www.cabq.gov/abq-data/>.
6. "Ashburn, VA." Weather Forecast & Reports. N.p., n.d. Web. 29 Nov. 2014. <http://www.wunderground.com/>.

1. Example of a TWO-COLUMN figure caption: (a) this is the format for referencing parts of a figure.