

# The Bee Corp's Method of Hive Strength Assessment for Pollination Effectiveness

**Ellie Symes**

Chief Executive Officer at the Bee Corp

[ellie.symes@thebeecorp.com](mailto:ellie.symes@thebeecorp.com)

**Joseph Cazier, Ph.D**

Appalachian State University

[cazierja@appstate.edu](mailto:cazierja@appstate.edu)

## Abstract

To support the goal of Apimondia's Bee Data Standardization Group to build a global standard for data related to bees and beekeeping The Bee Corp is sharing their internal standards to assess pollination effectiveness. Data standards presented in this report are used internally for measuring the effectiveness of beehives for crop pollination. The data provided includes output data that is shared with Bee Corp's users. *This is not a statement on data shared by the Bee Corp—this is simply the data standard used internally by the Bee Corp.*

*Key Words:* The Bee Corp, Honey Bees, Pollination Effectiveness, Data Standard

*Apimondia Scientific Commision:* Pollination and Bee Flora

## Introduction

Pollination is critical to the  $\sim \frac{3}{4}$  of agricultural crops (87 out of 115 primary food crops, Bauer and Wing, 2010) that need pollinated and the farmers and consumers dependant on those crops. In the U.S., 60% of revenue collected by beekeepers comes from pollination fees (Bond et al., 2014). However much of the contracted pollination process is opaque. Farmers don't always know what they are getting in terms of pollination effectiveness and beekeepers don't always get full value for the services they provide.

The Bee Corps mission is to develop technology to help beekeepers maximize pollination revenue per hive and growers optimize pollination costs and quality. By helping beekeepers measure and communicate the value of their bees, we aim to create a transparent marketplace that rewards excellent bees. To help in this process, the Bee Corp has developed a set of metrics to help farmers and beekeepers measure the *effectiveness* of pollination.

This article is about some of the metrics used by the Bee Corp to measure data and calculate outputs related to pollination effectiveness in the hopes that other groups learn from our experience and be able to share data efficiently among those that wish.

In this article, we begin with a review of the raw data collected, followed by the calculations we make with the raw data and then complete with a presentation of the proposed BeeXML standard for this system.

## **Raw Data Collected**

In this section we lay out and describe the raw data collected for readers to understand the goals and purpose of each measurement.

### Hive Strength or Frame Count

*How it's measured:* This is measured by the number of frames in an individual hive that are at least 75% covered by bees. This information can be collected three ways: a visual inspection of the hive, weighing the mass of the bees and converting bee mass to frame count, or through Bee Corp's Verifli technology which computes the number of bees and converts it to frame count.

*Data type:* Floating decimal

### Hive Location

*How it's measured:* This is measured using GPS technology, which provides latitude and longitude values in the decimal degrees format.

*Data type:* Text string

### Data tags

*How it's measured:* These tags allow qualitative information to be associated with individual frame strengths. Tags could associate people (example: beekeepers or growers), location names, or other information. A TagID allows for the information to be anonymous, but still grouped for analysis.

*Data type:* TagID and Name are both text strings

### Total number of hives rented

*How it's measured:* This is a sum of all the hives rented for pollination by a given grower.

*Data type:* Integer

### Total number of hives graded

*How it's measured:* This is a sum of all the hives with associated frame counts.

*Data type:* Integer

### Capture time

*How it's measured:* This is the date and time that a frame count is collected. The information is presented in YYYY-MM-DD HH:MM:SS format.

*Data type:* Text string

## **Computed Data**

In this section we lay out and describe the computed data calculated for readers to understand the goals and purpose of each calculation.

### Average Frame Strength

*How it's measured:* The average frame strength is the total sum of all the frame counts divided by the total number of hives in a set. The sets can be broken up by locations, beekeepers, growers, date, or other data tags.

*Data type:* Floating decimal

### Seasonal Average Frame Strength

*How it's measured:* This is computed by summing all of the frame strengths gathered in a single year for a single crop and dividing by the total number of hives graded in that year for that crop.

*Data type:* Floating decimal

### Distribution of Frame Counts

*How it's measured:* This is computed by summing the number of hives with a given frame count. Frame counts are rounded to the nearest integer.

*Data type:* List of floating decimals

### Frame Count Minimum and Maximum

*How it's measured:* This is computed by taking the lowest frame count (Minimum) and the highest frame count (Maximum) for a given set.

*Data type:* Floating decimal

### Percentage of Hives Graded

*How it's measured:* This is computed by taking the sum of hives graded and dividing by the total number of hives rented in a given set.

*Data type:* Floating decimal

## **Proposed XML Tags**

This section lays out the proposed tags for the XML standard for the human reader along with any relevant validation information and the rationale.

### Root Node: Hive Strength

The root node is the base node for this data. Since this is the data collected by the Bee Corp, to avoid confusion with other measures collected by others that have similar names, we suggest starting with the prefix BC to identify and credit the Bee Corp for this standard.

Since the goal of this assessment is to measure the strength of the hive, it seems logical to include that in the name. It also seems likely that this measurement will be done many times in many locations and would therefore need an ID. Recall from our (currently [draft](#)) tutorial for writing these articles that the data should be stored with tags and meta-data (data about data) should be stored as attributes. Therefore we recommend the root node have an ID attribute. The text of the ID can be whatever the user records, we just recommend it be unique to facilitate exchange.

Also note the recommendation for tag names is to use Title\_Case connected by an underscore.

Therefore we propose the following root tag.

```
<BC_Hive_Strength BCHS_ID = "insert ID here">
```

### Primary Child Node: Raw Data

Since there is both raw and calculated data in the assessment it seems likely that some users will want just the raw data and then do the calculations on their own to reduce transfer time, while other users may want both the raw and calculated fields for convenience or to just collect the summary data calculated for analysis.

To facilitate this exchange for both types of users, we create a child node for each group. Since this grouping is for organizational purposes, no data needs to be associated with the tag. Therefore we just need to make sure the tag is unique to make sure it is not confused with others. Hence we suggest BCHS\_Raw\_Data for the Bee Corp Hive Strength Assessment Raw Data. No attributes or validation are needed.

```
<BCHS_Raw_Data>
```

### Secondary Child Node: Hive Strength or Frame Count

Since the frame count will likely vary with the type of hive, we recommend an attribute to identify the type of hive.

<BCHS\_Frame\_Count Hive\_Type = “*insert type of hive here*”>

### Secondary Child Node: Hive Location

It is common to record GPS coordinates in either Degrees, Minutes, and Seconds (DMS) or Degrees and Decimal Minutes (DMM). Either can work and they can be converted, but it should be identified with a required attribute for clarity.

<BCHS\_Hive\_Location Degree\_Type = “*insert DMS or DMM here*”>

### Secondary Child Node: TagID

Note that a TagID is defined differently than an ID on the hive or a measurement ID. It can be that, but it can also be for a named location with several sharing a yard, or it can be for a beekeeper or something else. Therefore this is listed separately.

<BCHS\_TagID>

### Secondary Child Node: Total Number of Hives Rented

<BCHS\_Hives\_Rented>

### Secondary Child Node: Total Number of Hives Graded

<BCHS\_Hives\_Graded>

### Secondary Child Node: Capture Time

Since time may be recorded differently in different locations, especially with different time zones, we recommend a time zone attribute based on GMT.

<BCHS\_Time GMT = “*+/- GMT*”>

### Primary Child Node: Computed Data

Because this is computed over a group of data, as opposed to individually like the raw data, information is needed on which group the data pertains too to facilitate data sharing.

<BCHS\_Computed Grouping = “*insert group info here*”>

Secondary Child Node: Average Frame Strength

<BCHS\_Avg\_Fr\_Str>

Secondary Child Node: Seasonal Average Frame Strength

<BCHS\_Avg\_Sea\_Str>

Secondary Child Node: Distribution of Frame Counts

<BCHS\_HS\_Dist>

Secondary Child Node: Frame Count Minimum and Maximum

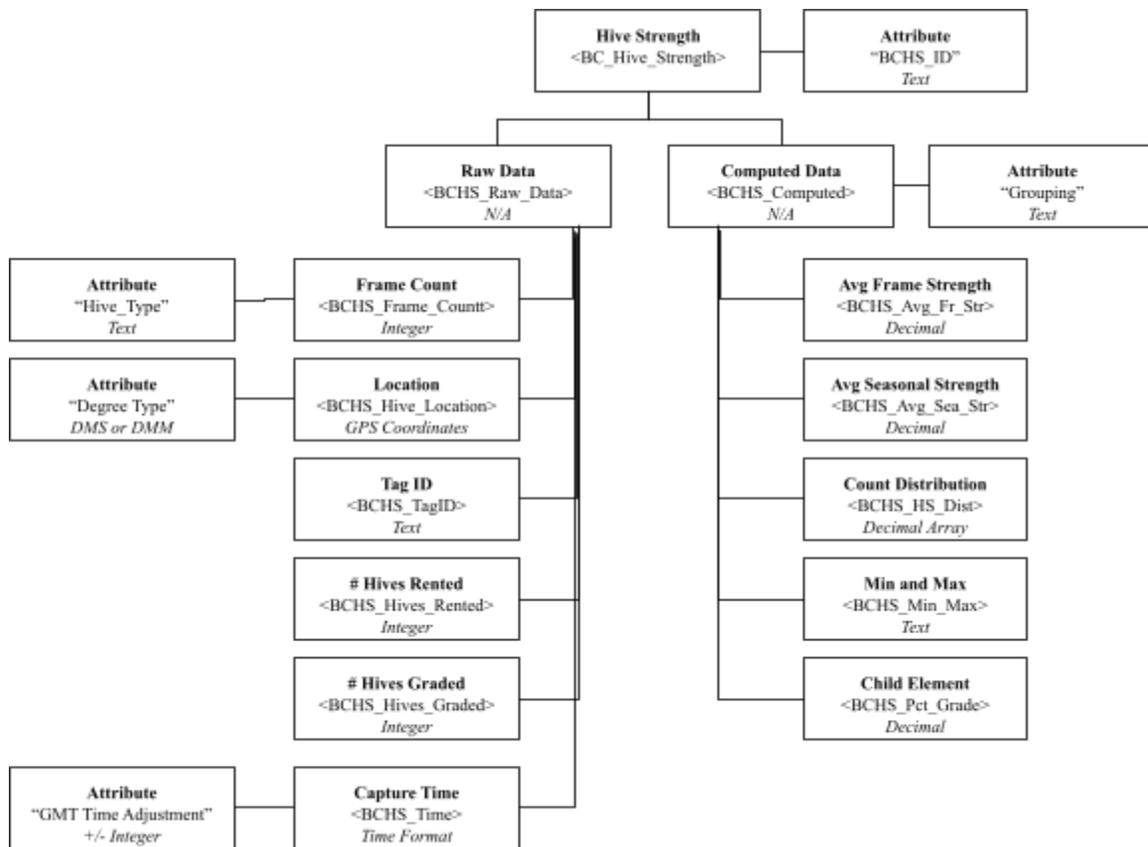
<BCHS\_Min\_Max>

Secondary Child Node: Percentage of Hives Graded

<BCHS\_Pct\_Grade>

**XML Tree**

This section provides a visualization of the BeeXML tree structure for quick reference.



## Example Data Document

The following is an example of the code and type of data that might be shared through a BeeXML Document. We include XML comments `<!-- comment -->` to explain the document to human readers. These are a best practice to use, but not strictly necessary for the technology to work. We will indent and italicize the comments for clarity. White space also added for readability.

```

<!-- File name loaded on Server will be 2020_Bee_Corp_Hive_Strength.xsd -->
<!-- Location will be in the directory http://beexml.appstate.edu/xml/Pollination_and_Bee_Flora
to match with the relevant scientific commission -->

<?xml version="1.0" encoding="UTF-8"?>

    <!-- This tells the system we are using XML and in standard 8-bit unicode characters -->

    <BC_Hive_Strength BCHS_ID = "North Yard Assessment"
        xmlns = "https://beexml.appstate.edu/Pollination_and_Bee_Flora"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="https://beexml.appstate.edu/xml/ Pollination_and_Bee_Flora.xsd">

        <!-- This header does a serval things

```

*First line Sets the root node and defines that ID for the assessment in the attribute*

*Second line creates a default namespace to be sorted by the Apimondia Scientific Commission to avoid overlapping names from different commissions.*

*Third line points to the World Wide Web (W3) consortium that defines the XML standard to give schema creation rules.*

*Fourth line points to the XML Schema document defining what the tags we created mean*  
-->

<BCHS\_Raw\_Data> <!-- Sets raw data child node -->

<BCHS\_Frame\_Count Hive\_Type = "Double Deep"> 7 </BCHS\_Frame\_Count>

<BCHS\_Location Degree\_Type = "DMS">

36° 12' 24.60" N -81° 40' 4.19" W <!-- Location of Boone, NC -->

</BCHS\_Location>

<BCHS\_TagID> 123456789 </BCHS\_TagID>

<BCHS\_Hives\_Rented> 250 </BCHS\_Hives\_Rented>

<BCHS\_Hives\_Graded> 100 </BCHS\_Hives\_Graded>

<BCHS\_Time GMT = "-5"> <!-- Adjust to East Coast Time in U.S. -->

2020-10-09 11:42:32

</BCHS\_Time>

</BCHS\_Raw\_Data>

<BCHS\_Computed Grouping = "Hives in North Yard for 2020 Pollination" >

<BCHS\_Avg\_Fr\_Str> 7.8 </BCHS\_Computed>

<BCHS\_Avg\_Sea\_Str> 8.3 </BCHS\_Avg\_Sea\_Str>

<BCHS\_HS\_Dist>

*<!-- An example of how a distribution could be communicated with the first number the count and the second the number in that count. Ended at 10 frames as this is for illustration purposes only -->*

0:20 1:10 2:5 3:55 4:75 5:100 6:150 7:350 8:500 9:400 10:85

</BCHS\_HS\_Dist>

<BCHS\_Min\_Max> 3:10 </BCHS\_Min\_Max>

<BCHS\_Pct\_Grade> 25.7 </BCHS\_Pct\_Grade>

</BCHS\_Computed>

</BCHS\_Hive\_Strength>

*<!-- Can add as many additional data points as you like below -->*

Note how the tags open and close in the order they were created, almost likely poetry it rhymes. An XML parser can read this code and exchange data into the format needed for the new data storage system or leave it stored as an XML file.

## **BeeXML Schema**

An XML Schema is the dictionary that defines the data and data type. It can also add restrictions such as labeling something required or specifying an order. Below is a proposed schema document for the Bee Corp Hive Strength Assessment.

<?xml version = "1.0" encoding = "UTF-8"?>

*<!-- This header information defines the xml version used and the unicode character format -->*

<xs:schema xmlns:xs = "http://www.w3.org/2001/XMLSchema"  
targetNamespace = "http://beexml.appstate.edu/xml/Pollination\_and\_Bee\_Flora"  
xmlns = "http://beexml.appstate.edu/Pollination\_and\_Bee\_Flora">

*<!-- This links to the W3 consortium schema website that hosts information useful for data validation and formats. Next it sets up a namespace to avoid confusion among similar tags used for other bee data by linking it to one of seven Apimondia Scientific Commissions -->*

*<!-- Define elements here -->*

*<!-- Root Element -->*

```
<xs:element name = "BC_Hive_Strength" type = "xs:string"/>
```

```
    <xs:attribute name = "ID" type = "xs:string"/>
```

```
<!-- Raw Data Elements -->
```

```
<xs:element name = "BCHS_Raw_Data" type = "xs:string"/>
```

```
    <xs:element name = "BCHS_Frame_Count" type = "xs:integer"/>
```

```
        <xs:attribute = "Hive_Type" type = "xs:string"/>
```

```
    <!-- Note that for now we are listing Hive_Type as a string. Strongly suggest that  
    as soon as one of our teams works to standardize names for common hive types  
    that this attribute is updated to match -->
```

```
    <xs:element name = "BCHS_Location" type = "xs:string"/>
```

```
        <!-- The code below sets up the two ways to record GIS data, it is  
        restricted to these two enumerated values for quality control -->
```

```
        <xs:attribute name = "Degree_Type" use = "required">
```

```
            <xs:complexType>
```

```
                <xs:restriction base = "xs:string">
```

```
                    <xs:enumeration value = "DMS"/>
```

```
                    <xs:enumeration value = "DMM"/>
```

```
                </xs:restriction>
```

```
            </xs:complexType>
```

```
        </xs:attribute>
```

```
    <xs:element name = "BCHS_TagID" type = "xs:string"/>
```

```
    <xs:element name = "BCHS_Hives_Rented" type = "xs:integer"/>
```

```
    <xs:element name = "BCHS_Hives_Graded" type = "xs:integer"/>
```

```
    <xs:element name = "BCHS_Time" type = "xs:dateTime"/>
```

```
        <xs:attribute name = "GMT" type = "xs:integer"/>
```

```
    <!-- Note that we could have included the UTC or GMT indicator in the time  
    value itself with a timecode. However, since we are dealing with nature both  
    absolute time and relative time matter and will likely be used for different types
```

*of analysis, therefore an attribute seemed to make more sense so the data can more easily be stored separately and merged when needed. -->*

*<!-- Computed Data Elements -->*

```
<xs:element name = "BCHS_Computed"/>
    <xs:attribute name = "Grouping" type = "xs:string"/>
<xs:element name = "BCHS_Avg_Fr_Str" type = "xs:decimal"/>
<xs:element name = "BCHS_Avg_Sea_Str" type = "xs:decimal"/>
<xs:element name = "BCHS_HS_Dist" type = "xs:string"/>
<xs:element name = "BCHS_Min_Max" type = "xs:string"/>
<xs:element name = "BCHS_Pct_Grade" type = "xs:decimal"/>

</xs:schema>
```

## **Acknowledgments**

The editors are especially grateful for Ellie Symes of the Bee Corp for volunteering to be one of the first to volunteer to share the structure of their data for this purpose.

## **References**

Bauer, D., & Wing, I. (2010). Economic Consequences of Pollinator Declines: A Synthesis. *Agricultural and Resource Economics Review*, 39(3), 368-383. doi:10.1017/S1068280500007371

Jennifer Bond, Kristy Platter, and Hunt Kevin. (2014). *Fruit and Tree Nuts Outlook: Economic Insight. U.S. Pollination Services Market*. U.S. Department of Agriculture, Economic Research Service. [https://www.ers.usda.gov/webdocs/outlooks/37059/49130\\_fts-357.pdf?v=4060](https://www.ers.usda.gov/webdocs/outlooks/37059/49130_fts-357.pdf?v=4060)

## Appendix A - XML Schema Document

```
<?xml version = "1.0" encoding = "UTF-8"?>

<!-- This header information defines the xml version used and the unicode character format -->

<xs:schema xmlns:xs = "http://www.w3.org/2001/XMLSchema"
  targetNamespace = "http://beexml.appstate.edu/xml/Pollination_and_Bee_Flora"
  xmlns = "http://beexml.appstate.edu/Pollination_and_Bee_Flora">

  <!-- This links to the W3 consortium schema website that hosts information useful for
  data validation and formats. Next it sets up a namespace to avoid confusion among
  similar tags used for other bee data by linking it to one of seven Apimondia Scientific
  Commissions -->

  <!-- Define elements here -->

  <!-- Root Element -->

    <xs:element name = "BC_Hive_Strength" type = "xs:string"/>

    <xs:attribute name = "ID" type = "xs:string"/>

  <!-- Raw Data Elements -->

    <xs:element name = "BCHS_Raw_Data" type = "xs:string"/>

    <xs:element name = "BCHS_Frame_Count" type = "xs:integer"/>

    <xs:attribute = "Hive_Type" type = "xs:string"/>

    <!-- Note that for now we are listing Hive_Type as a string. Strongly suggest that
    as soon as one of our teams works to standardize names for common hive types
    that this attribute is updated to match -->

    <xs:element name = "BCHS_Location" type = "xs:string"/>

    <!-- The code below sets up the two ways to record GIS data, it is
    restricted to these two enumerated values for quality control -->

    <xs:attribute name = "Degree_Type" use = "required">
      <xs:complexType>
        <xs:restriction base = "xs:string">
          <xs:enumeration value = "DMS"/>
          <xs:enumeration value = "DMM"/>
        </xs:restriction>
      </xs:complexType>
    </xs:attribute>
  </xs:schema>
```

```

        </xs:restriction>
        </xs:complexType>
    </xs:attribute>

    <xs:element name = "BCHS_TagID" type = "xs:string"/>

    <xs:element name = "BCHS_Hives_Rented" type = "xs:integer"/>

    <xs:element name = "BCHS_Hives_Graded" type = "xs:integer"/>

    <xs:element name = "BCHS_Time" type = "xs:dateTime"/>

        <xs:attribute name = "GMT" type = "xs:integer"/>

    <!-- Note that we could have included the UTC or GMT indicator in the time
    value itself with a timecode. However, since we are dealing with nature both
    absolute time and relative time matter and will likely be used for different types
    of analysis, therefore an attribute seemed to make more sense so the data can
    more easily be stored separately and merged when needed. -->

<!-- Computed Data Elements →

    <xs:element name = "BCHS_Computed"/>

        <xs:attribute name = "Grouping" type = "xs:string"/>

    <xs:element name = "BCHS_Avg_Fr_Str" type = "xs:decimal"/>

    <xs:element name = "BCHS_Avg_Sea_Str" type = "xs:decimal"/>

    <xs:element name = "BCHS_HS_Dist" type = "xs:string"/>

    <xs:element name = "BCHS_Min_Max" type = "xs:string"/>

    <xs:element name = "BCHS_Pct_Grade" type = "xs:decimal"/>

</xs:schema>

```