Operational Applications of the VHI for Real-Time Crop Yield Modeling and Assessments at USDA’s World Agricultural Outlook Board

Eric Luebehusen
Meteorologist
USDA/OCE/WAOB
Background & History...

1999; NESDIS-SAB: Operational Meteorologist using satellite imagery for DVORAK, VAAC, Flooding, Special Events.

1999-2005; NOAA-NWS-CPC: JAWF (@ USDA) covering weather in SOA, SEA, AUS.
The World Agricultural Outlook Board (WAOB) serves as USDA's focal point for economic intelligence and the commodity outlook for U.S. and world agriculture. WAOB coordinates, reviews, and approves the monthly World Agricultural Supply and Demand Estimates (WASDE) report as well as long-term Agricultural Baseline Projections. It is also the focal point for analyzing weather-related impacts on agriculture through the Office of the Chief Meteorologist. In addition, the Board is responsible for coordinating the World Agricultural Outlook Forum, USDA's oldest and largest gathering.

**2005-Current; WAOB:** Primary responsibility is to provide crop-yield weather intelligence in support of the WASDE.
WASDE: Market-sensitive document released by the 12th of every month detailing the latest US and global crop situational outlook.

“Trading Places”

WHEAT: The outlook for 2022/23 U.S. wheat this month is for increased supplies, unchanged domestic use and exports, and higher stocks. Supplies are raised on higher production with all wheat production projected at 1,737 million bushels, up 8 million from last month. NASS raised winter wheat production to 1,182 million bushels as increases for Soft Red Winter and White Winter more than offset a reduction for Hard Red Winter. The all wheat yield is 46.9 bushels per acre, up 0.3 bushels from last month. Projected 2022/23 ending stocks are raised 8 million bushels to 627 million, still down 4 percent from 2021/22. The projected 2022/23 season-average farm price is unchanged at $10.75 per bushel, compared to $7.70 for 2021/22.

The global wheat outlook for 2022/23 is for lower supplies, reduced consumption, fractionally lower trade, and slightly lower ending stocks. Supplies are decreased by 1.7 million tons to 1,052.8 million as lower India production more than offsets an increase for Russia. India’s production is lowered 2.5 million tons to 106.0 million as extreme temperatures in March and April reduced yields during grain fill. Russia’s production is raised 1.0 million tons to 81.0 million with all of the increase in winter wheat on generally favorable weather conditions to date. Projected 2022/23 world consumption is reduced 1.5 million tons to 786.0 million mainly on lower feed and residual use for India and less food, seed, and industrial use for Sri Lanka and Argentina.

Projected 2022/23 global trade is decreased 0.3 million tons to 204.6 million as lower exports from India are not completely offset by higher exports from Russia and Uzbekistan. India’s exports are reduced 2.0 million tons to 6.5 million as the government intends to restrict exports to some destinations to ensure sufficient domestic supplies. Russia’s exports are raised 1.0 million tons to 40.0 million, which would be the second largest on record. Russia’s supplies are projected higher for 2022/23 and its export prices are more competitive than most other exporters. Projected 2022/23 world ending stocks are lowered 0.2 million tons to 266.9 million, a six-year low.

COARSE GRAINS: This month’s 2022/23 U.S. corn outlook is for larger beginning stocks, slightly higher use, and increased ending stocks. Corn area and yield forecasts are unchanged. USDA will release its Acreage report on June 30, which will provide survey-based indications of planted and harvested area. Beginning stocks are up 45 million bushels mostly reflecting a forecast decline in exports for 2021/22. Exports are lowered 50 million
• Weekly weather assessments and write ups for the WWCB
• Monthly or as-needed weather briefings detailing latest crop-weather information
• Operational yield modeling updated weekly in support of the WASDE
Over the past several years, the ability to quantify crop stages using Growing Degree Days (GDDs) and match them with the VHI and Weather (Wx) data have led to an ever-expanding crop-yield modeling effort at USDA-WAOB.

I am currently running 126 operational international crop-country yield regression forecast suites (7 forecasts each workbook).

I have also added 43 for the US.
It all starts with data. The main source of WMO weather data for the WAOB is through a MOU with NOAA/NWS/CPC; this is the backbone of our work. Data is maintained in house at USDA.
10,000+ stations
Dates back to 1980
120,000,000 lines of data and counting
Additional SQL database server for domestic data
Global dataset is supplemented with national-level data
It’s a LOT of daily Wx data!

<table>
<thead>
<tr>
<th>FID</th>
<th>Shape</th>
<th>WMO</th>
<th>LATITUDE</th>
<th>LONGITUDE</th>
<th>PRECIP</th>
<th>TAVG</th>
<th>TMAX</th>
<th>TMIN</th>
<th>NPRECIP</th>
<th>HTAVG</th>
<th>TANOM</th>
<th>PHP</th>
<th>PANOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>133</td>
<td>Point</td>
<td>01001</td>
<td>70.93</td>
<td>-8.67</td>
<td>2</td>
<td>-2</td>
<td>2</td>
<td>-6</td>
<td>10</td>
<td>-4</td>
<td>2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>134</td>
<td>Point</td>
<td>01002</td>
<td>80.02</td>
<td>16.25</td>
<td>0</td>
<td>-10</td>
<td>-8</td>
<td>-13</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>135</td>
<td>Point</td>
<td>01003</td>
<td>77</td>
<td>15.5</td>
<td>0</td>
<td>-9</td>
<td>-5</td>
<td>-12</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>136</td>
<td>Point</td>
<td>01006</td>
<td>78.25</td>
<td>22.83</td>
<td>0</td>
<td>-11</td>
<td>-7</td>
<td>-15</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>137</td>
<td>Point</td>
<td>01007</td>
<td>78.92</td>
<td>11.93</td>
<td>2</td>
<td>-10</td>
<td>-6</td>
<td>-14</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>138</td>
<td>Point</td>
<td>01008</td>
<td>78.25</td>
<td>15.47</td>
<td>0</td>
<td>-10</td>
<td>-6</td>
<td>-16</td>
<td>2</td>
<td>-11</td>
<td>1</td>
<td>0</td>
<td>-2</td>
</tr>
<tr>
<td>139</td>
<td>Point</td>
<td>01009</td>
<td>60.67</td>
<td>20.85</td>
<td>0</td>
<td>-11</td>
<td>-5</td>
<td>-13</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>140</td>
<td>Point</td>
<td>01010</td>
<td>69.3</td>
<td>16.15</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>-2</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-17</td>
</tr>
<tr>
<td>141</td>
<td>Point</td>
<td>01011</td>
<td>80.08</td>
<td>31.38</td>
<td>0</td>
<td>-10</td>
<td>-2</td>
<td>-15</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>142</td>
<td>Point</td>
<td>01015</td>
<td>69.6</td>
<td>17.83</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>-2</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>143</td>
<td>Point</td>
<td>01017</td>
<td>69.35</td>
<td>18.08</td>
<td>0</td>
<td>-6</td>
<td>-5</td>
<td>-11</td>
<td>13</td>
<td>1</td>
<td>-9</td>
<td>0</td>
<td>-13</td>
</tr>
<tr>
<td>144</td>
<td>Point</td>
<td>01023</td>
<td>69.05</td>
<td>18.55</td>
<td>3</td>
<td>0</td>
<td>7</td>
<td>-9</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>-5</td>
</tr>
<tr>
<td>145</td>
<td>Point</td>
<td>01025</td>
<td>69.68</td>
<td>18.92</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>-4</td>
<td>14</td>
<td>1</td>
<td>0</td>
<td>29</td>
<td>-10</td>
</tr>
<tr>
<td>146</td>
<td>Point</td>
<td>01026</td>
<td>69.65</td>
<td>16.94</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>-4</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>150</td>
<td>Point</td>
<td>01028</td>
<td>74.52</td>
<td>19.02</td>
<td>3</td>
<td>-4</td>
<td>-1</td>
<td>-9</td>
<td>6</td>
<td>-5</td>
<td>1</td>
<td>50</td>
<td>-3</td>
</tr>
<tr>
<td>151</td>
<td>Point</td>
<td>01029</td>
<td>70.7</td>
<td>30.06</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>-5</td>
<td>8</td>
<td>-1</td>
<td>2</td>
<td>0</td>
<td>-8</td>
</tr>
<tr>
<td>153</td>
<td>Point</td>
<td>01092</td>
<td>70.37</td>
<td>31.1</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>-4</td>
<td>9</td>
<td>-1</td>
<td>2</td>
<td>44</td>
<td>-5</td>
</tr>
<tr>
<td>154</td>
<td>Point</td>
<td>01098</td>
<td>65.2</td>
<td>11</td>
<td>0</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>12</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>-12</td>
</tr>
<tr>
<td>155</td>
<td>Point</td>
<td>01102</td>
<td>65.7</td>
<td>11.85</td>
<td>15</td>
<td>4</td>
<td>8</td>
<td>-4</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>156</td>
<td>Point</td>
<td>01112</td>
<td>65.47</td>
<td>12.22</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>-1</td>
<td>13</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>-13</td>
</tr>
<tr>
<td>157</td>
<td>Point</td>
<td>01115</td>
<td>66.75</td>
<td>12.48</td>
<td>14</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>15</td>
<td>4</td>
<td>0</td>
<td>93</td>
<td>-3</td>
</tr>
<tr>
<td>158</td>
<td>Point</td>
<td>01116</td>
<td>65.97</td>
<td>12.47</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>-3</td>
<td>-9999</td>
<td>3</td>
<td>1</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>159</td>
<td>Point</td>
<td>01121</td>
<td>66.37</td>
<td>12.62</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>0</td>
<td>19</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>-19</td>
</tr>
<tr>
<td>160</td>
<td>Point</td>
<td>01122</td>
<td>65.78</td>
<td>13.22</td>
<td>3</td>
<td>3</td>
<td>9</td>
<td>7</td>
<td>-9999</td>
<td>2</td>
<td>1</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>161</td>
<td>Point</td>
<td>01139</td>
<td>67.68</td>
<td>12.68</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>-1</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>162</td>
<td>Point</td>
<td>01141</td>
<td>65.15</td>
<td>13.62</td>
<td>0</td>
<td>3</td>
<td>8</td>
<td>-3</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
<tr>
<td>163</td>
<td>Point</td>
<td>01147</td>
<td>65.52</td>
<td>14.02</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>-12</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
<td>-9999</td>
</tr>
</tbody>
</table>

Agricultural Weather Assessments
World Agricultural Outlook Board
Much of the data analysis and visualization developmental work has been done in Excel.

*Software developed, updated, and maintained in-house.

Agricultural Weather Assessments
World Agricultural Outlook Board
Using established GDD *crop-stage cutoffs*, real-time estimates of *stages of development for crops across the globe* within Excel are possible using our built-in Wx DB links.
When I began this effort back in **2015**, impact assessments were subjective.
Meanwhile, we began downloading the 16 km VHI data back in 2007 and archiving it for use in our Weather Presentations in support of the WASDE.

This slide is from a 2010 presentation I gave to the Chairman of the WAOB regarding Europe Wheat.
In 2015, Dr. Felix Kogan and his staff began distributing the 4 km VHI geotiff data (back built to 2005); the difference was visually striking.
However, we were limited to calculating statistics within GIS, a slow and cumbersome process.
A significant breakthrough: In June of 2017, the Admin-00 and Admin-01 VHI ascii data became available to download. I began to work with the data that summer; this would launch the crop yield modeling effort.
The final piece: At WAOB’s request, NESDIS added a 
croplands-masked VHI dataset in 2019, working 
with the weather group to test and implement a 
consensus global croplands filter that was applied 
to the entire dataset.
The Process...
Operational note: **Downloading the ascii files** was initially time consuming and labor intensive. However, a vital component to our operations is **WGET** download utility...
WGET batch jobs download thousands of Admin 00 and Admin 01 ascii files automatically to our server, allowing us to focus on data analysis.
At last count, every week we are downloading over 4,000 VHI ascii files covering more than 35 countries to support our expanding international crop yield modeling effort.
Water Deficit Yield Response Factors (FAO) depict the key yield determinant phases of development; for grains it’s typically flower, while oilseeds are generally a bit later (fill).
Note: It is very important at WAOB to use crop stages versus week numbers.

i.e. In 2022, winter rapeseed over England developed 1-2 weeks ahead of average and well ahead of last year, illustrating the utility of GDDs.
Crop Stage “Week Number” charts over the past 30 years illustrate the year-to-year variability across many growing areas.

~ June 1st

~ May 25th

~ April 27th
Crop Stage “Week Number” charts over the past 30 years illustrate the year-to-year variability across many growing areas.

- September 7th
- August 17th
- July 27th
Within Excel, **Crop-Stage-Specific VHI & Weather** are extracted for yield forecasting using regression analysis.

*Using crop stages versus static week numbers ensures year-to-year consistency & allows easy testing of different scenarios.*
VHI Yield Regression....
Each Regression Workbook contains 7 different equations, with the VHI used in 5 of them.
Each Regression Workbook contains 7 different equations, with the VHI used in 5 of them.
The first order of business is to go into the VHI Forecast worksheet within the modeling workbook and select an expected “high impact” VHI crop stage wrt to yields to see if there are any disconnects or other issues.
In this case, a cursory view of Ukraine Corn Yield and VHI @ Silk shows a good correlation from 2007 onward, but prior to 2007 the fit does not appear to be strong. We can easily adjust Start and End Years for the modeling.
Each stage is tested, starting with Tassel.
Next up, test Silk stage VHI.
Blister, slightly better stats than Silk.
Dough, Stats start to slip.
Dent,
Worse stats still.
There are other Excel tricks, bells, and whistles, but the gist: All stages are tested, and for Ukraine Corn, the best VHI stats are during **Blister**
The VHI stage is set to Blister and rerun to set that as the VHI forecast parameter moving forward.
VHI data sources can be adjusted as well.
It is possible to pick other admins, the entire country, or other user-defined averages (CSV files).
The VHI ascii data is VITAL to our operations!!
Wx Regression....
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to **test every Wx scenario** (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to **test every Wx scenario** (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
A macro allows me to test every Wx scenario (Weather Start, Stop combo) and have the results populate in a 4x4 array.
For Ukraine Corn, the best results are Silk→Blister.
Other Regression....
Ukraine Corn Regression Stats

2007-2020

Fcast-VHI
\[ R^2 = 0.87 \]
\[ Adj \, R^2 = 0.85 \]
Excellent

Fcast-Wx
\[ R^2 = 0.71 \]
\[ Adj \, R^2 = 0.62 \]
Good

Fcast-VHI+Wx
\[ R^2 = 0.9 \]
\[ Adj \, R^2 = 0.85 \]
Very Good

Fcast-VHI+Wx Intcpt
\[ R^2 = 0.9 \]
\[ Adj \, R^2 = 0.86 \]
Very Good

Fcast-VHI+Wx2 Intcpt
\[ R^2 = 0.9 \]
\[ Adj \, R^2 = 0.84 \]
Very Good

Stand Alone
\[ R^2 = 0.8 \]
\[ Adj \, R^2 = 0.78 \]
Excellent

6.84
3.47

Significance-F
(Objective Assessment)
Excellent \( F < 0.0001 \)
Very Good \( F < 0.001 \)
Good \( F < 0.05 \)
Fair \( F < 0.1 \)
Poor \( F \geq 0.1 \)
The “Combo” regression (3, 4, & 5) are often the highest performers, but R² is weighed against Adj R², Sig-F test, tStats, and Pearson coefficients.
The VHI with no Trend and Yield Trend are run in conjunction with the other guidance.
Output....
All forecasts are summarized in chart form; this is from September 2021 for Ukraine Corn. These forecasts ended up being very accurate.

Final Yield: 7.68
A “Stats” chart is produced; this contains statistics for each equation to assess the forecasts’ utility and whether there are things that need correcting. It also guides which equation(s) I should lean on. I use this chart A LOT!!
A “Hindcast” is also created using the RESIDUAL OUTPUT of eq’ns 1, 2, 3, and either 4 or 5 (Incpt Shift; user-selects). This helps assess potential equation biases.
The GDD crop stage and Weekly VHI summarize crop development wrt to the VHI. This has proven very persuasive.
Other behind-the-scenes charts are produced as well, including the **Stage-Specific VHI & Yield** for the past 30 years (if available).
A summary of all forecast data broken down by region (FSU shown here) is published and sent to the different commodity groups. This contains a great deal of information; many columns were added as a direct result of questions I have heard during the Commodity meetings.
Results....
What follows is a quick overview of the forecast performance covering the past 4 years.
Good news: The Peak $R^2$ Regression should beat Trend, and it does.
And hopefully, my suggestion (SuggestionE) beats the Regression (it does)!
Model ( ■ ) Forecast Rank by Crop
1. Sunflowers
2. Corn
3. Soybeans
4. Wheat
5. Rapeseed
6. Barley
7. Cotton

2018-21 Median Forecast Error - By Crop

- Wheat: 5.7%, SuggestionE 5.7%, Peak R² Fcst 6.9%
- Oilseed, Rapeseed: 7.9%, SuggestionE 7.9%, Peak R² Fcst 7.8%
- Oilseed, Soybean: 5.3%, SuggestionE 5.3%, Peak R² Fcst 6.1%
- Oilseed, Sunflowerseed: 4.8%, SuggestionE 4.8%, Peak R² Fcst 4.9%
- Barley: 6.2%, SuggestionE 6.2%, Peak R² Fcst 8.0%
- Corn: 3.8%, SuggestionE 3.8%, Peak R² Fcst 5.4%
- Cotton: 10.5%, SuggestionE 10.5%, Peak R² Fcst 19.8%
Conclusions....
We have been using the VHI in some form operationally since 2007.
The VHI ascii data and croplands masked ascii data were game changers for USDA-WAOB operations.