**Baseliner 4 Instructions**

**(Updated 8/3/2016)**

**1. Overview**

The Baseliner (version 4) software is an interactive tool for processing sap flux data from thermal dissipation probes (TDP; Granier 1985, 1987). It is designed to provide a graphical user interface for (1) simple, automated data cleaning, (2) visual inspection and manual data cleaning, and (3) estimation of sap flux from raw TDP data. Baseliner is freely available, open source software. For additional information, please refer to Oishi et al. (2016). We request that you cite this reference in any publications that utilize this software.

**2. Thermal Dissipation Probe Operation**

TDPs measure the temperature difference (dT) between heated and unheated probes. dT is inversely related to sap flux (*F*) as a function of the maximum temperature differential between probes when flux is zero (dTmax). Thus, *F* can be related to dT as:

*F* = a \* ((dTmax – dT) / dT) ^ b = a \* ((dTmax / dT) - 1) ^ b Eq. 1

where a and b are empirical coefficients. Since a and b have been shown to vary among some species, Baseliner does not estimate *F*, but rather *K*, a dimensionless flow index as defined by Granier (1985):

*K*= (dTmax –dT)/dT = (dTmax / dT) – 1 Eq. 2

Thus, the software offers maximum user flexibility.

dTmax varies among sensors and changes over time due to tree water status, ambient temperature fluctuations, and power supplied to the heating element, among other factors (Lu et al. 2004). Additionally, “noise” in sensor data can lead to misinterpretation of the true data signal. For these reasons, no universal rule or algorithm exists for identifying dTmax. This software takes a hybrid approach to dTmax estimation by first identifying points in time where flow is likely zero, based on biophysical conditions, then allowing the user to visually inspect and modify those points. After identifying individual dTmax “anchor points”, a continuous dTmax, or “baseline” is linearly interpolated for the entire time series.

Baseliner 4 offers two methods for automatically selecting dTmax anchor points. The default option identifies the maximum dT on each night were data are available. For this approach, night is defined as the time from midnight to 7:00 AM, a window of time when the atmospheric demand for water is low and recharge of depleted stem water is likely to have ceased. This approach is based on the assumption that sap flux declines to zero every night; however, since nocturnal sap flux has been observed in a variety of tree species and ecosystems (e.g. Dailey & Phillips 2006; Dawson et al., 2007; Oishi et al., 2008), this assumption may not be ideal for all data. Nevertheless, this method can be used if radiation and vapor pressure deficit (VPD) data are not available and may provide a reasonable initial estimate of dTmax.

The second approach is based on identifying times when sap flux is likely zero, described in Oishi et al. (2008). The conditions that must be met for likely zero-flow are:

1. Nighttime hours (defined by near-zero radiation),
2. dT values within a specified time period (default = 2-hours) with a low coefficient of variation (i.e. stable dT), and
3. near-zero VPD for a defined period (below a specified value for a set amount of time).

For each night when all 3 conditions are met, only the final dT point in that night is selected. For all dT points meeting conditions 1 & 2 and meeting conditions 1, 2, & 3, alternative symbols are displayed, identifying other plausible dTmax points.

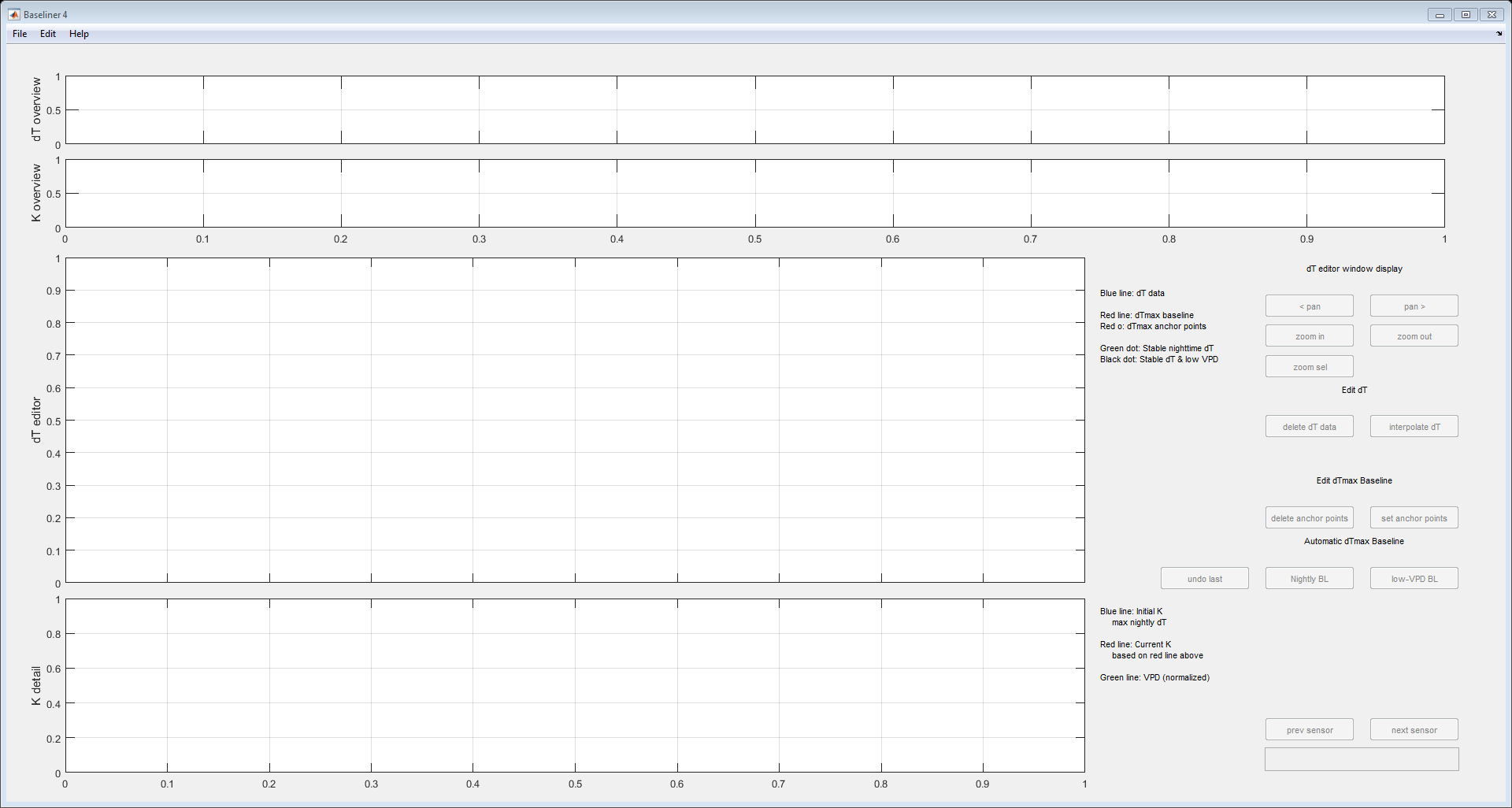
**3. General overview of Baseliner 4’s file structure**

1. The raw data input file is a .csv file created by the user prior to starting Baseliner. This file is read by Baseliner, but not modified.
2. The “Project” file is an .xml file that Baseliner creates to track changes to the raw data file. It allows for the user to close and reopen an existing project and continue to work with the data. The file is readable in a text editor program, but is likely not useful.
3. The exported “K” data are the processed data using Baseliner. Units are arbitrary and should be used with the appropriate empirical relationship for a given species (e.g. *F* = 118.99x10-6 *K* 1.231 where F is sap flux density (m3 m-2 s-1); Granier 1987). Columns follow the raw input file structure.
4. The exported “K error” data are the estimated range of K associated with the variability of hand-selecting dTmax anchor points. This procedure runs Monte-Carlo simulations of the K estimates based on randomly selected dTmax points, varying around the selected dTmax anchor points within a normal distribution with a standard deviation of 1-hour (independent of time step). The output from this file is the standard deviation of exported K. Columns follow the raw input file.

**4. Interface window**

The interface window includes four plots and a series of buttons, allowing the user to visualize and edit the sap flux data.

* The “dT Overview” plot (top plot) shows the complete time series of the raw sap flux data for the selected sensor (starting at the first dT column provided in the data). The blue line represents the data, the red line represents the current zero-flow “baseline”, and the light blue box identifies the selected region of the data displayed in the third subplot.
* The “K Overview” plot (second from top) displays the converted sap flux based on the current zero-flow baseline for the entire time series. The blue line represents the converted data”, and the light blue box identifies the selected region of the data displayed in the fourth subplot.
* The “dT Editor” plot (third from top) shows the range of dT data within the light blue box in the full dT plot. This plot can be used as a graphical interface to edit raw data and the dTmax baseline
* The “K Detail” plot (bottom plot) shows the range of converted K-data within the light blue box in the full K-value plot. The blue line represents estimated sap flux based on the original, default dTmax. The red dashed line is sap flux based on the current, working dTmax line. The green line represents VPD (rescaled on a normalized scale from 0 to 1).

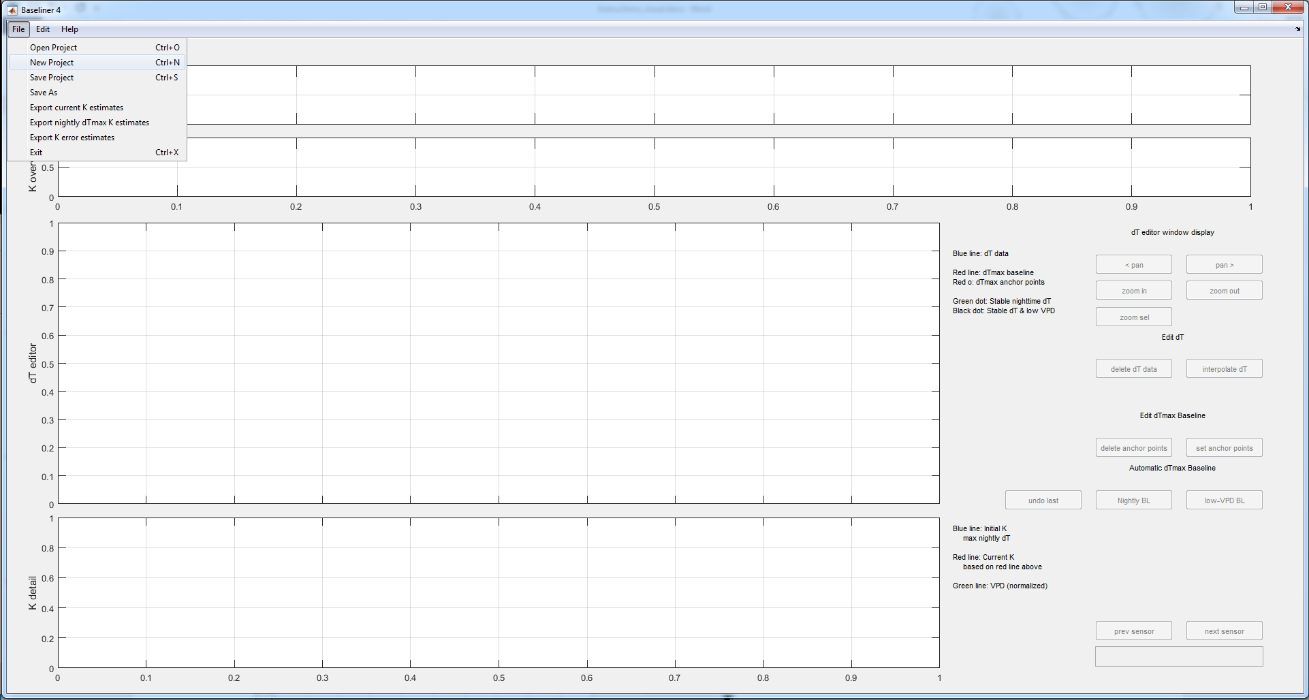


**5. Processing Sap Flux Data in Baseliner**

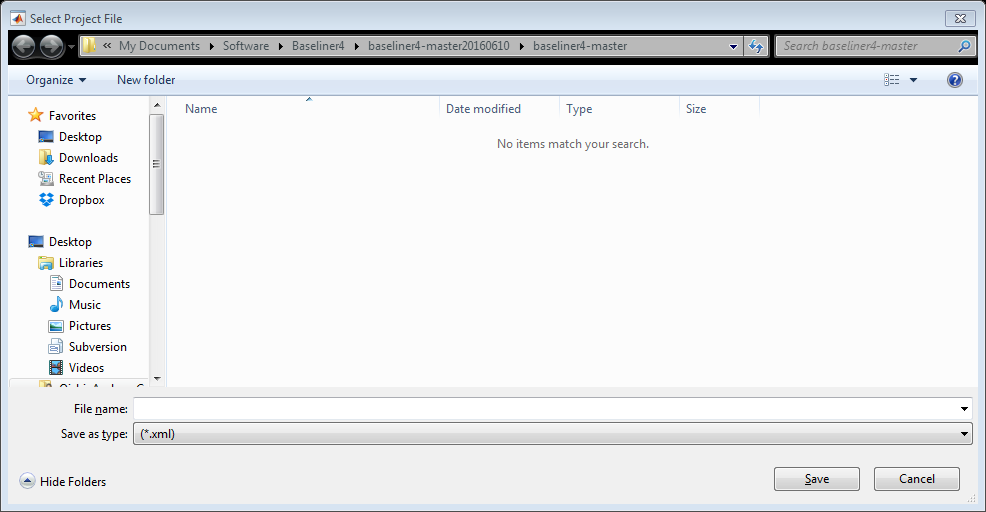
In order to view and process sap flux data, raw data are loaded and a “Project” file is created. The project accesses raw data, but does not alter this file. The Project file tracks changes made to the raw data, selected dTmax anchor points, and other metadata. The Project file can be saved and reopened to resume or reanalyze data.

**5.1 Starting a new Project**

Select the “File” tab, then the “New Project” option

A pop-up window will allow you to name the file associated with your new Project and designate its location. The Project information will be saved as an .xml file. Baseliner won’t alter your raw data and will export converted data into a .csv file.



You’ll also be prompted to select your input data file. **Input Data**

Data should be structured in a continuous time series with a regular time-step interval. There are not minimum or maximum limits to the time span of data. However, the sequence of days must increase, so we recommend limiting data to a single calendar year, or if data do span from December to January of the following year, modifying day of year to extend beyond 365. Any gaps in data (e.g. due to sensor outages) need to be filled with not-a-number (NaN) or an arbitrary number outside a specified range (e.g. -6999; further details on defining this range below). The minimum necessary data are dT from each sensor pair. This program uses two other meteorological variables to determine conditions conducive to zero-flow, vapor pressure deficit (VPD) and photosynthetically active radiation (PAR). The program also uses time of day to limit automatically-selected dTmax points earlier than 10:00 AM. In instances where VPD is not available, a placeholder variable, either zero or not-a-number (NaN) should be used. A zero value will allow for any point to be an automatically-selected potential zero-flow point, whereas NaN will prevent an automatically-selected point from being selected at this time. In instances where PAR is not available,

The data must be organized in the following columns:

1: Plot ID (for user reference, not used in processing)

2: Year (for user reference, not used in processing)

3: Day of year (continuous, monotonically increasing integers)

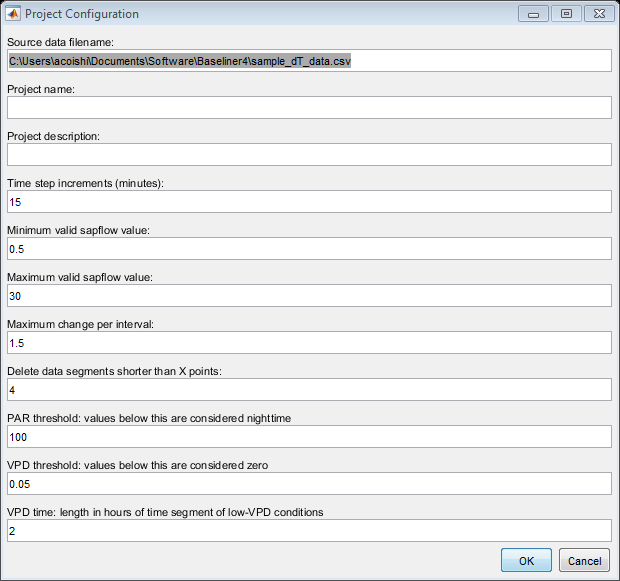
4: Time of day (format HHMM, such such that 1:45 AM = 145 and 3:00 PM = 1500)

5: Vapor pressure deficit (kPa, however any units will work, user defines threshold value)

6: Photosynthetically active radiation (any units will work, user defines day/night threshold)

6+n: dT data, Sensor set #n (either degrees C or mV)

You will then be prompted to designate some key parameters for your data set in the Project Configuration window.



|  |  |
| --- | --- |
| ***Field name*** | ***Description*** |
| Source data filename: | Name of dT input file (selected in previous window) |
| Project name: | Descriptive name of the current project (for user reference, not required) |
| Project description: | Additional description of current project (for user reference, not required) |
| Time step increments (minutes): | Time between each data point. 15 = 15 minutes. 60 = 1 hour. |
| Minimum valid sapflow value: | Filters data that are below acceptable range. Any dT data less than this value will be deleted. If you’re not sure what the minimum value is, you can set it to zero. |
| Maximum valid sapflow value: | Filters data that are above acceptable range. Any dT data greater than this value will be deleted. If you’re not sure what the maximum value is, you can set it to a very high number (e.g. 10000). |
| Maximum change per interval: | Filters data spikes. Any value with an absolute change from the previous time step that is greater than this value will be deleted. If you’re not sure what a reasonable value is, you can set it to a very high number (e.g. 10000). |
| Delete data segments shorter than X points: | Filters unusable, short data sequences. Sets minimum threshold for length of consecutive, valid dT points. If you’re not sure what a reasonable value is, you can set it to zero. |
| PAR threshold for nighttime | Threshold for determining nighttime hours. Note that due to PAR/radiation sensor calibration/drift issues or light pollution, sensors may read some nominal value greater than zero during the night. |
| VPD threshold: values below this are considered zero | Threshold for low-VPD conditions below which transpiration is expected to cease. Note that due to temperature/RH sensor calibration/drift issues, sensors may read some nominal value greater than zero during saturated conditions. |
| VPD time: length in hours of time segment of low-VPD conditions | Time threshold for low-VPD conditions below which transpiration is expected to cease. |

**5.2 Loading a previously started project**

Once you’ve started a project, you can open it again at any time to either finish data processing of all sensors, or re-edit previously processed sensors.

1. Click on the “File” tab at the top of the window
2. Select “Open Project”
3. Use the “Select Project File” window to identify a file location the existing file that you would like to open.

**5.3 Data QA/QC**

Once the data has been loaded into the interface window, the user has several options for QA/QC of the dT data and the subsequent selection of dTmax anchor points.

**5.3.1 Data visualization**

The editor window will initially display the first column of sap flux data from the raw file. Using the “prev sensor” and “next sensor” buttons, you can change which sensor you are editing. The current sensor number is displayed in the box in the lower right hand corner of the window. Once you’ve reached the last of your sensors, selecting “next sensor” will loop back to Sensor #1.

The entire time series of dT data will appear in the “dT overview” panel at the top of the editor window. Data within the shaded blue box is expanded in the “dT editor” panel in the center of the window. The “K overview” panel (second from the top) displays the calculated K, based on the default approach for dTmax anchor point selection (blue line) and K based on any modifications that have been made to the dTmax anchor points (dashed red line). The expanded view of the K data within the highlighted blue box is displayed in the “K detail” panel at the bottom of the window. The “K detail” panel also displays VPD (green line), which has been rescaled to arbitrary units to correspond with maximum K-values). Since sap flux is typically well-correlated with VPD (assuming non-limiting soil moisture and minimal time lags due to stem water capacitance), comparing converted K-values and VPD is a useful approach for checking data quality and dTmax anchor point selection.

You can zoom in or out using the “zoom” buttons on the right hand side of the window. You can also zoom in to a particular section of data by holding down your left mouse button and dragging to select a box. This feature will work for any of the four sub-plots. If you select a section in one of the upper two subplots, it will automatically zoom to that section. If you select a section in one of the lower two subplots, you will then need to click the “zoom sel” button. You can use the two “pan” buttons to shift the viewing area to the left or right.

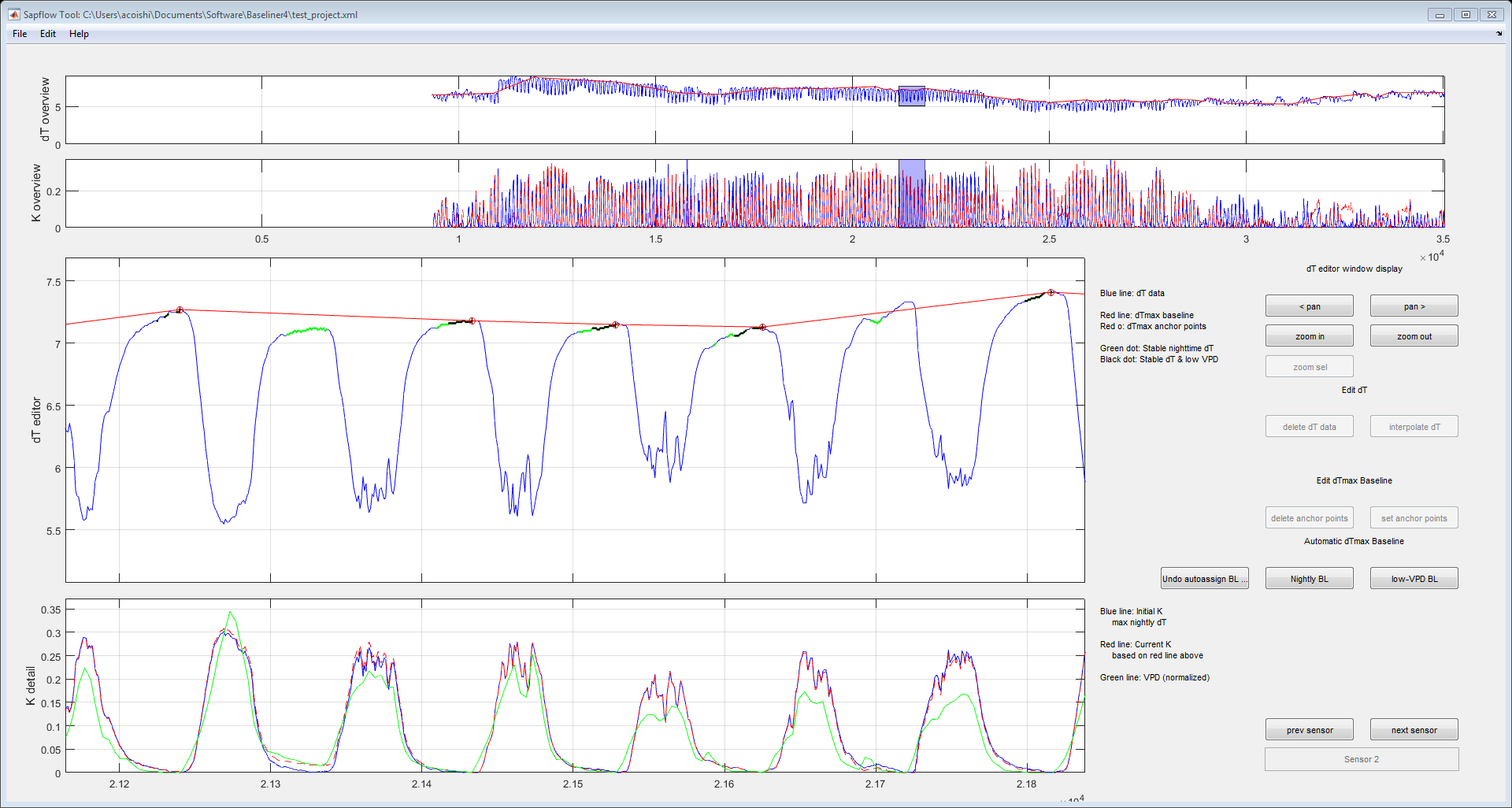






**5.3.2 Automatic Baseline Options**

The default approach for baseline estimation is using the maximum dT value each day prior to 07:00 AM. Selecting the “auto BL” function will select new baseline anchor points based on the criteria you specified when creating this Project. In this case, the program identified nighttime dT that was stable (green dots in main dT plot), VPD was not sufficiently low during most of these times, so only one dTmax anchor point was selected in this section of data.

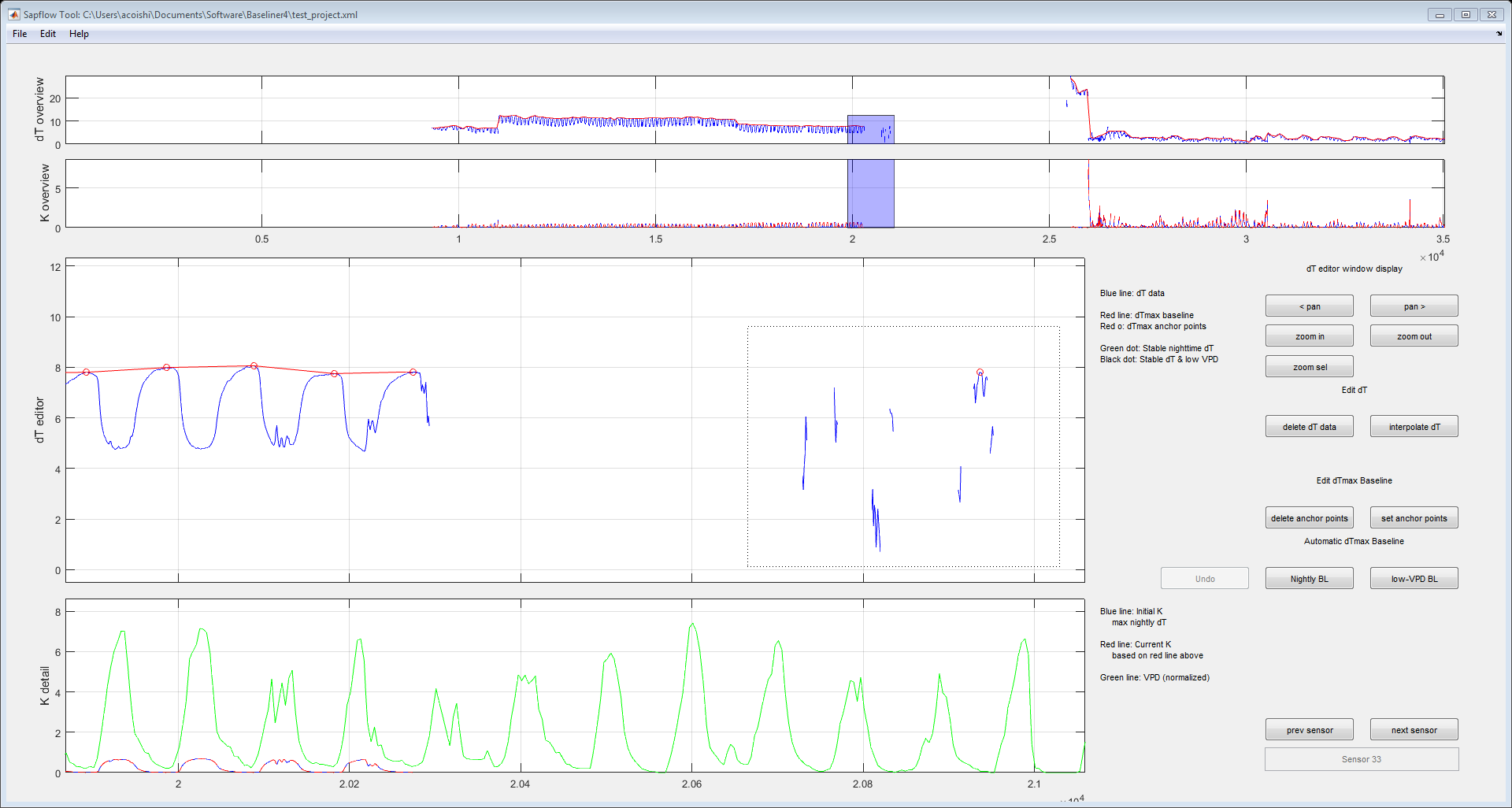


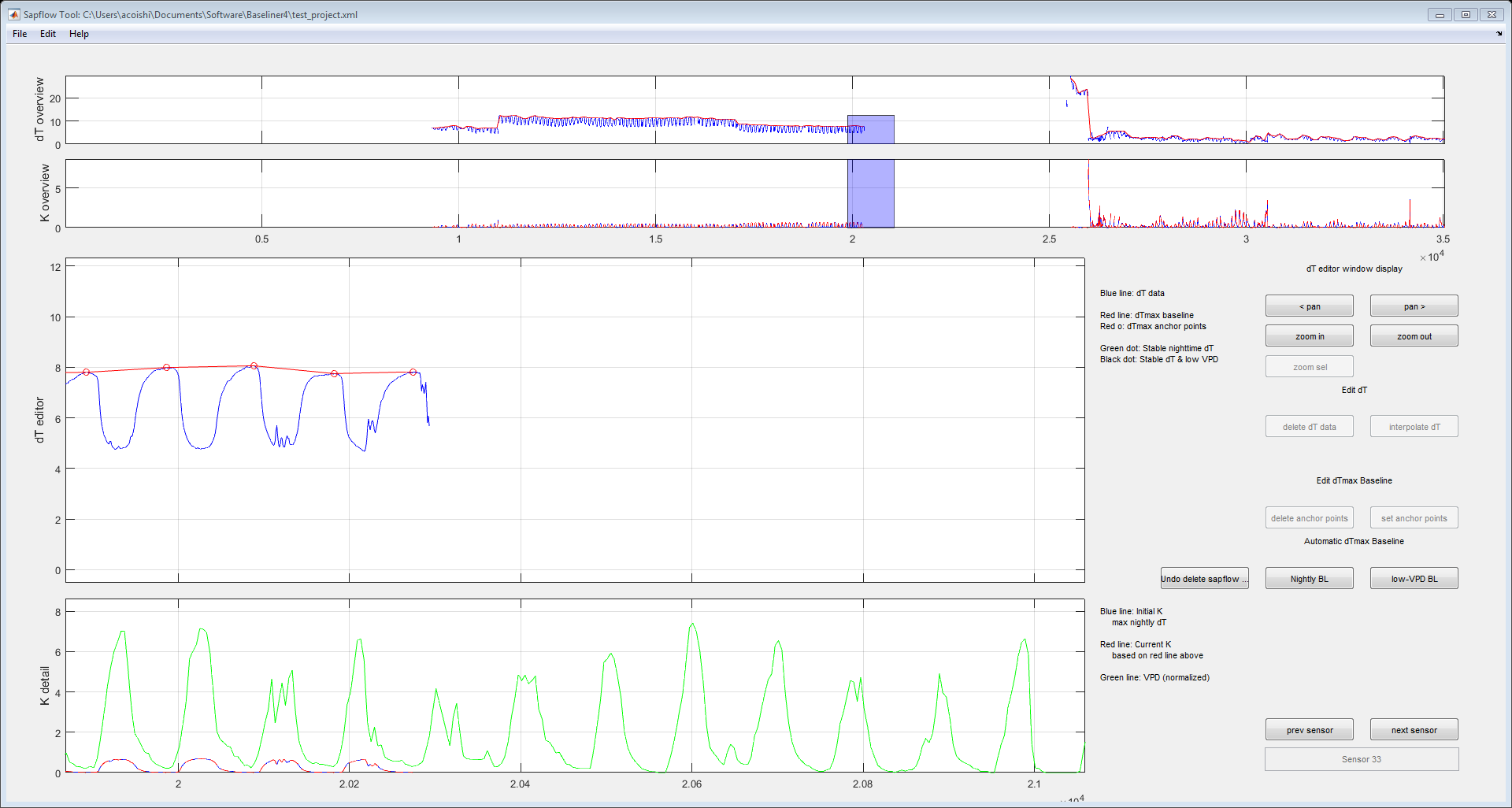
Also note the difference between the default baseline and this new baseline. For the first half of the time series in the lower K plot, the blue and red lines are very similar for much of the time.

Selecting the “Nightly BL” button will recalculate dTmax points using the default approach. Note that by selecting this option, you will lose any changes to dTmax points that you have made for the current sensor. (Any changes to dT values will be preserved.)

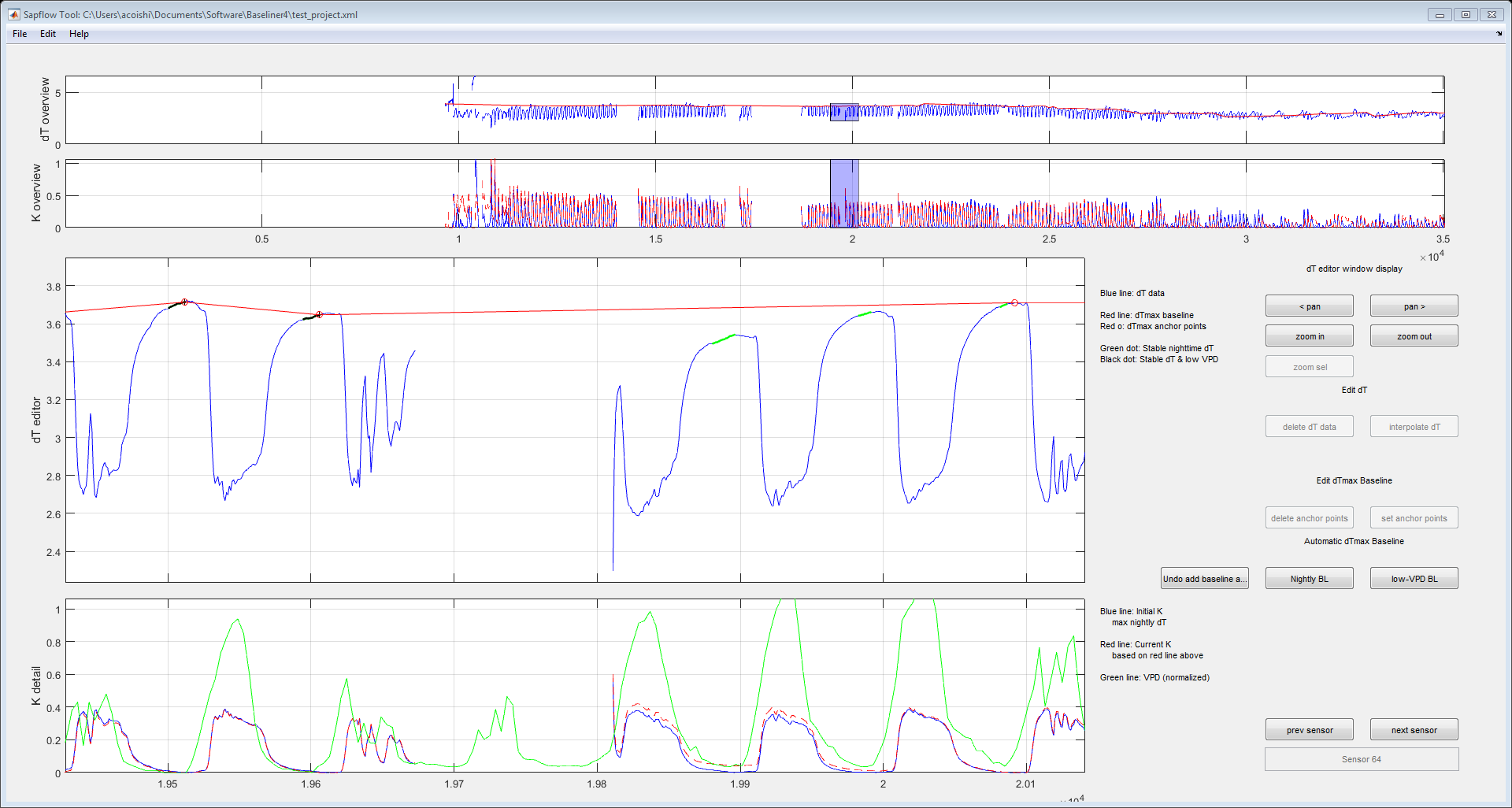
**5.3.3 Deleting dT data**

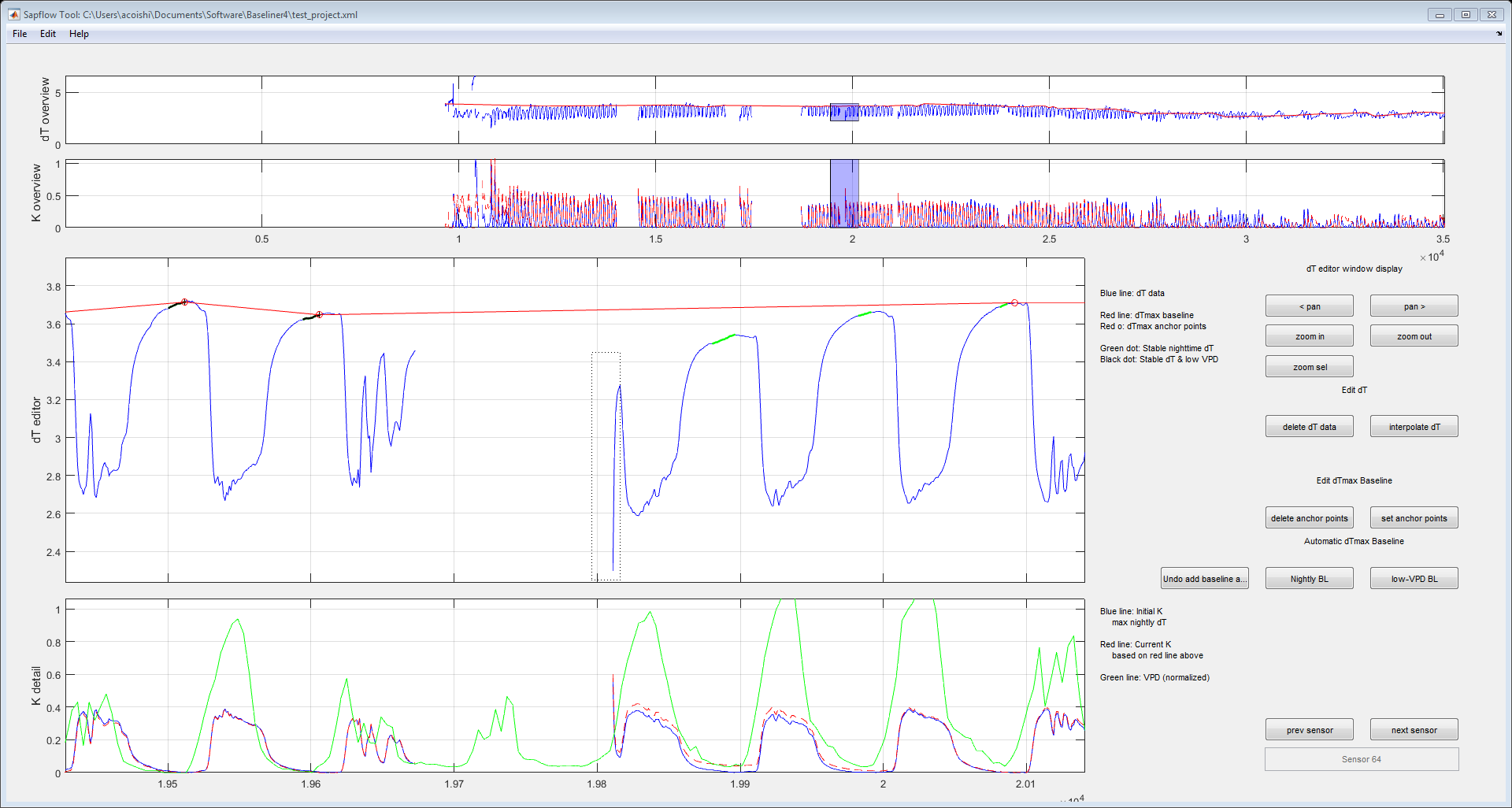
The Baseline program will clean out-of-range data and delete small fragments (pre-defined in the Project Configuration window). However, you will probably also need to remove additional unreliable data. In the following example, several short fragments of erratic data occur within a large data gap. These are likely output from a broken sensor. In order to fix this problem, use the mouse to select a box around the problematic data and either select the “delete dT data” button or press the “D” key on the keyboard.

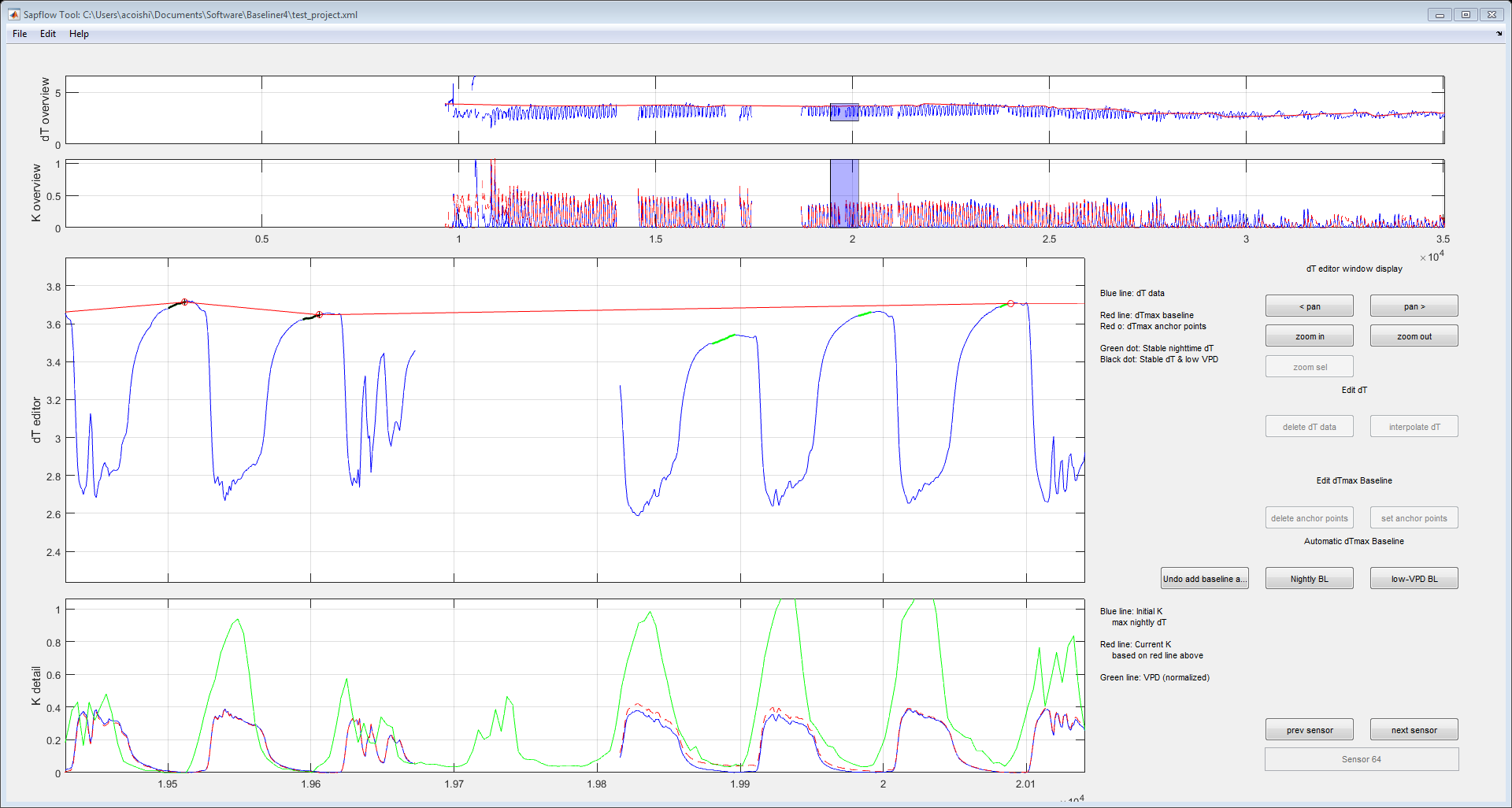




Data gaps are typically created when a sensor is not working properly or when the power to the heater has changed. In these instances, data around a gap is often unreliable. Since heater/sensor adjustments typically happen during the day, but dT anchor points occur at night, data between the anchor point and the break will still be converted and K-values may be inappropriately forced to zero or exaggerated.



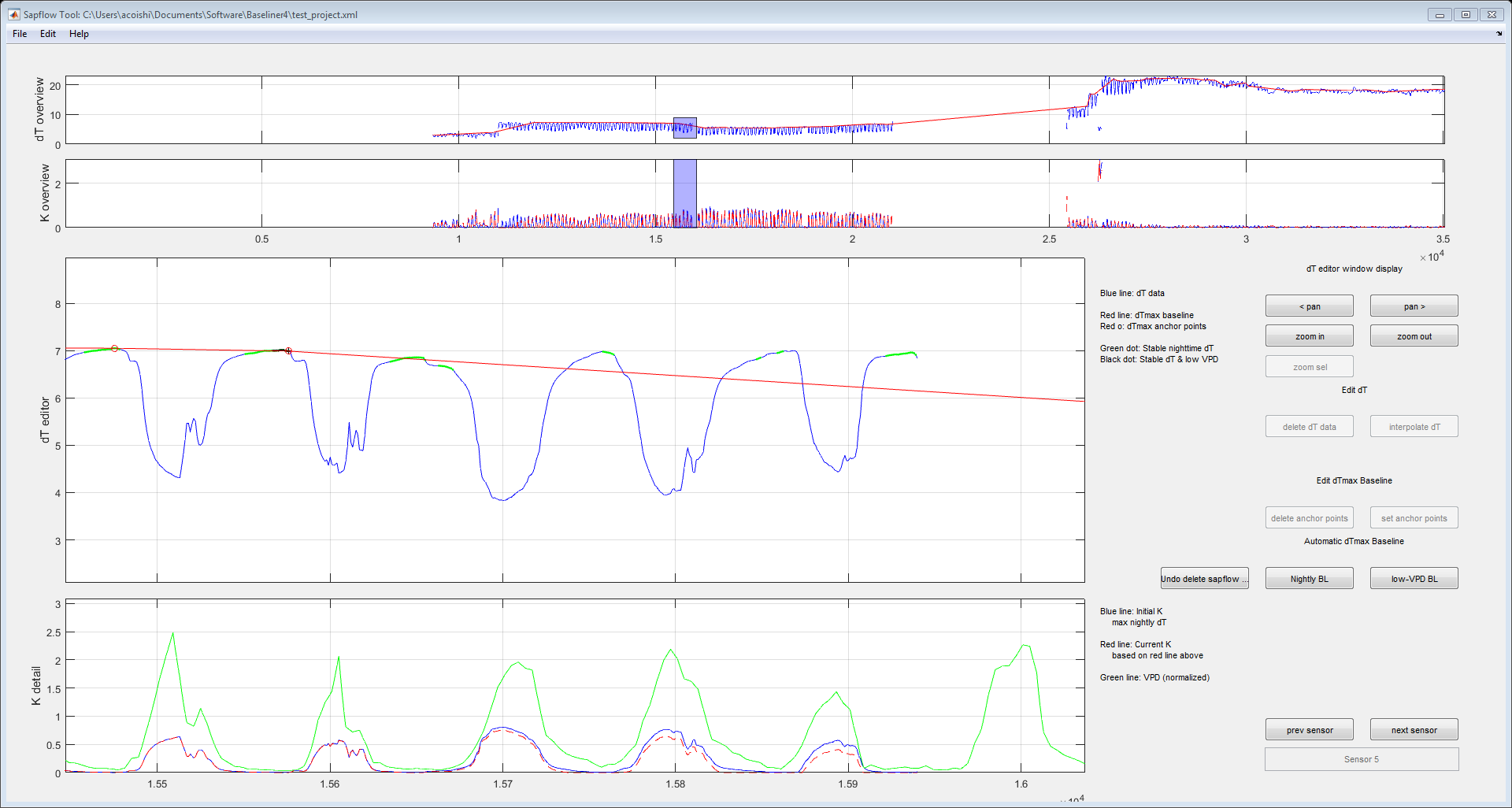


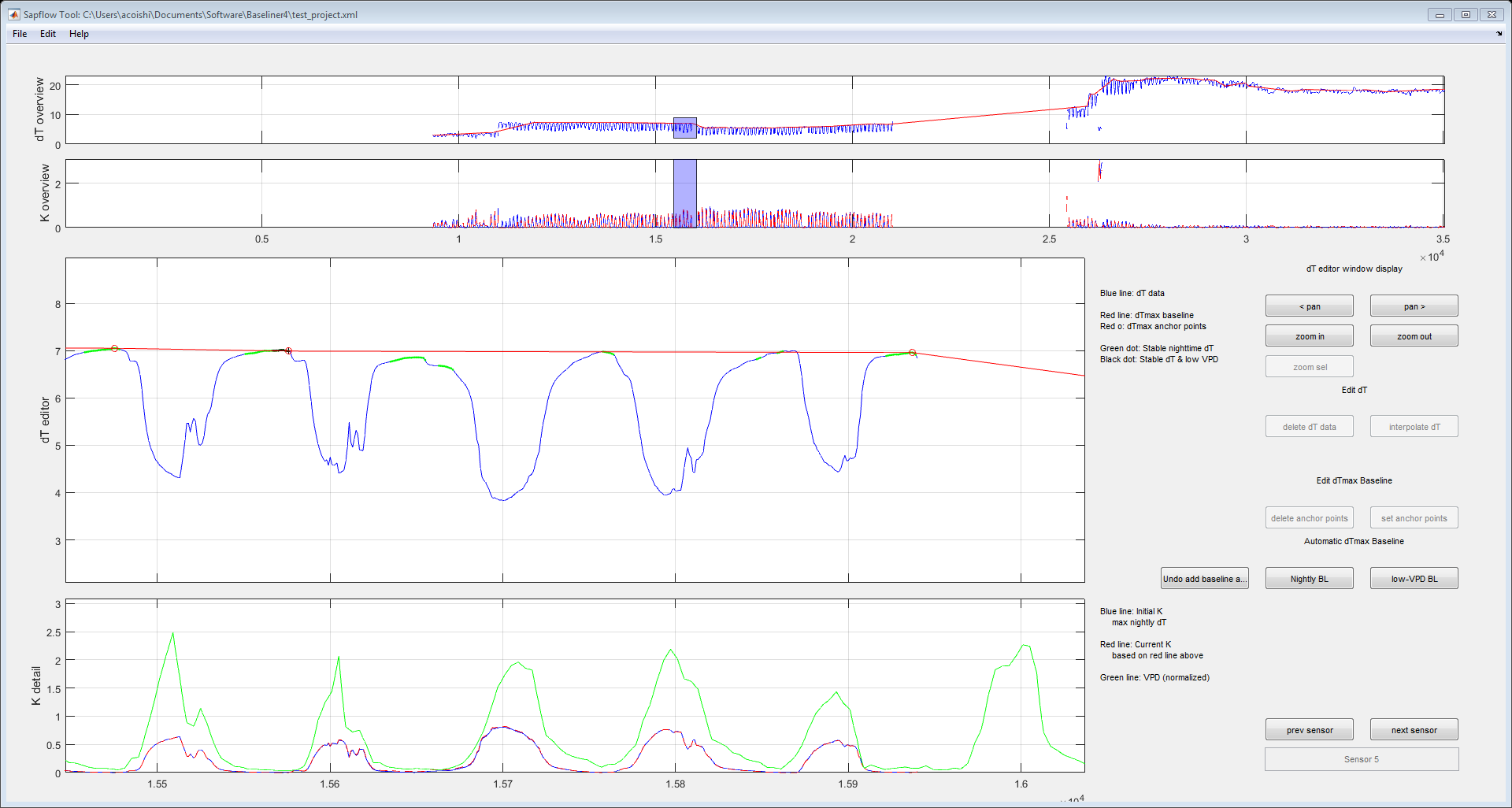


In this instance, the data before the gap look reasonable. The baseline is relatively stable before and after the gap and the converted K-values are similar to the previous days and appear to correspond with the VPD data. On the other hand, the dT data after the gap start off low (likely due to a power interruption and the heater warming back up), leading to a very high K-value that is out of sync with VPD. In this case, the first few data points following the gap should be deleted. A more conservative approach, would be to delete all of the data in the green box.

**5.3.4 Manually Adding dTmax Anchor Points**

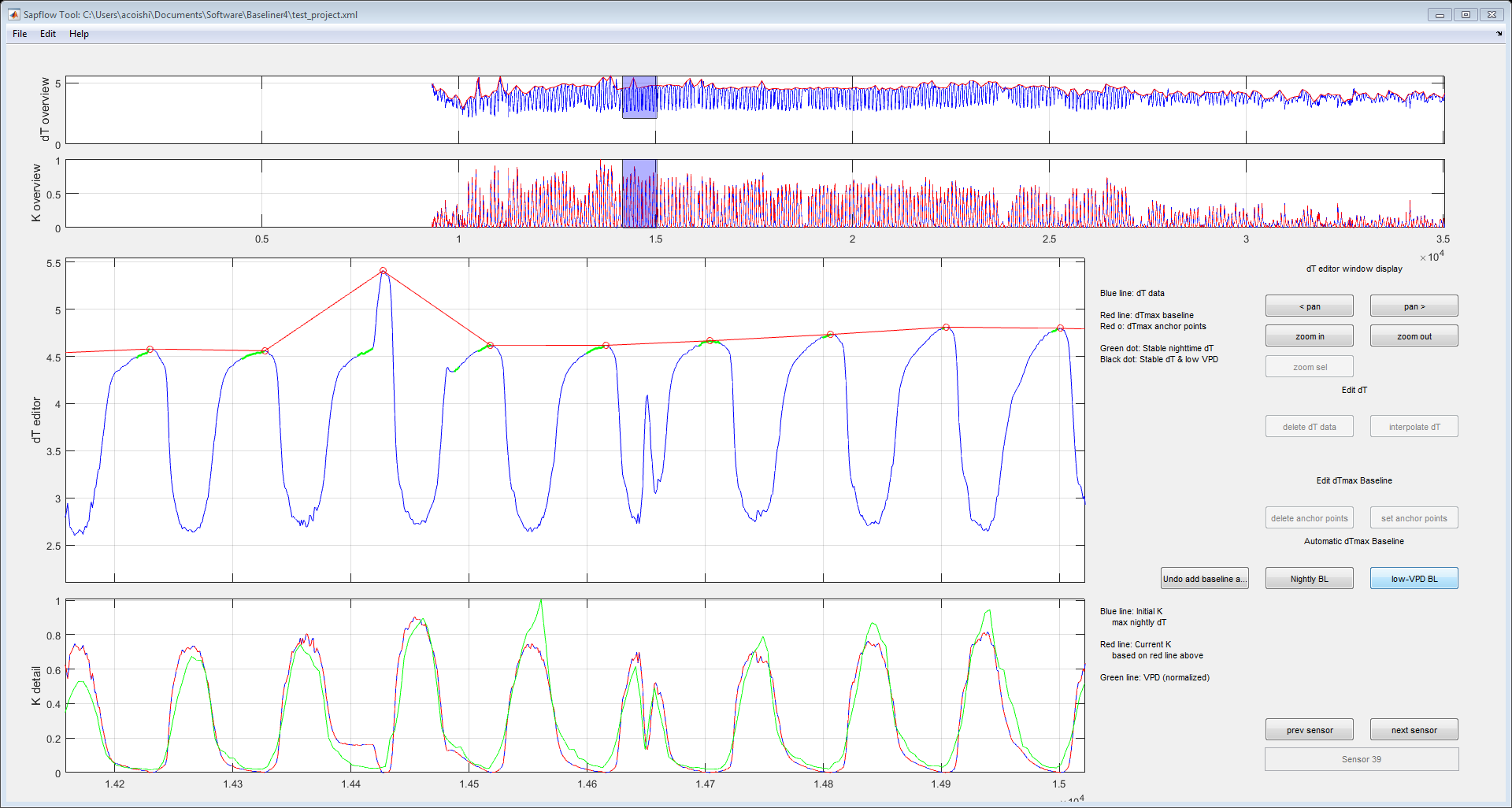
We can add a new baseline anchor by simply left-clicking at the designated point on the main dT plot. The K-value will automatically update. In this instance, since nighttime VPD was not below the threshold, an anchor point was not automatically set. Since the baseline (red) declines below the dT data, if we decided not to add a point on this night, we would likely need to delete dT data between the second anchor point and the data gap in the dT Editor window. In this instance, we could make the decision that VPD is sufficiently low and stable and add an anchor point on the final night before sensor failure.

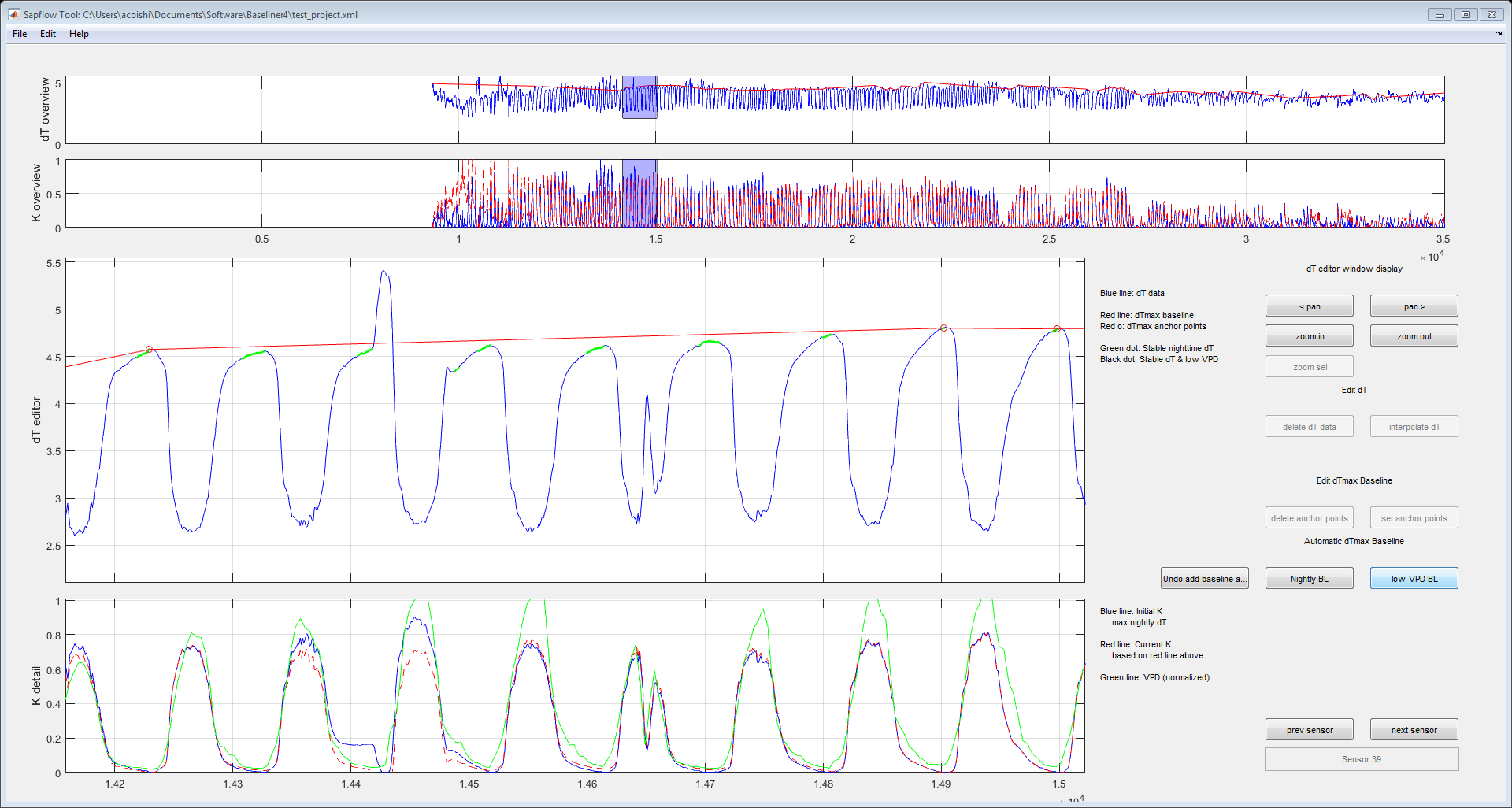




**5.3.5 Removing dTmax Anchor Points**

In some instances, you may see late-night/early-morning peaks in dT. These unstable peaks may deviate from the best estimate of the dTmax baseline. Assuming peak nightly dT is the baseline anchor point will lead to an overestimation of sap flux, particularly at night (see blue line on lowest subplot). Since Baseliner converts and dT data above the baseline to a K-value of zero, in this instance, no additional data adjustments are necessary.





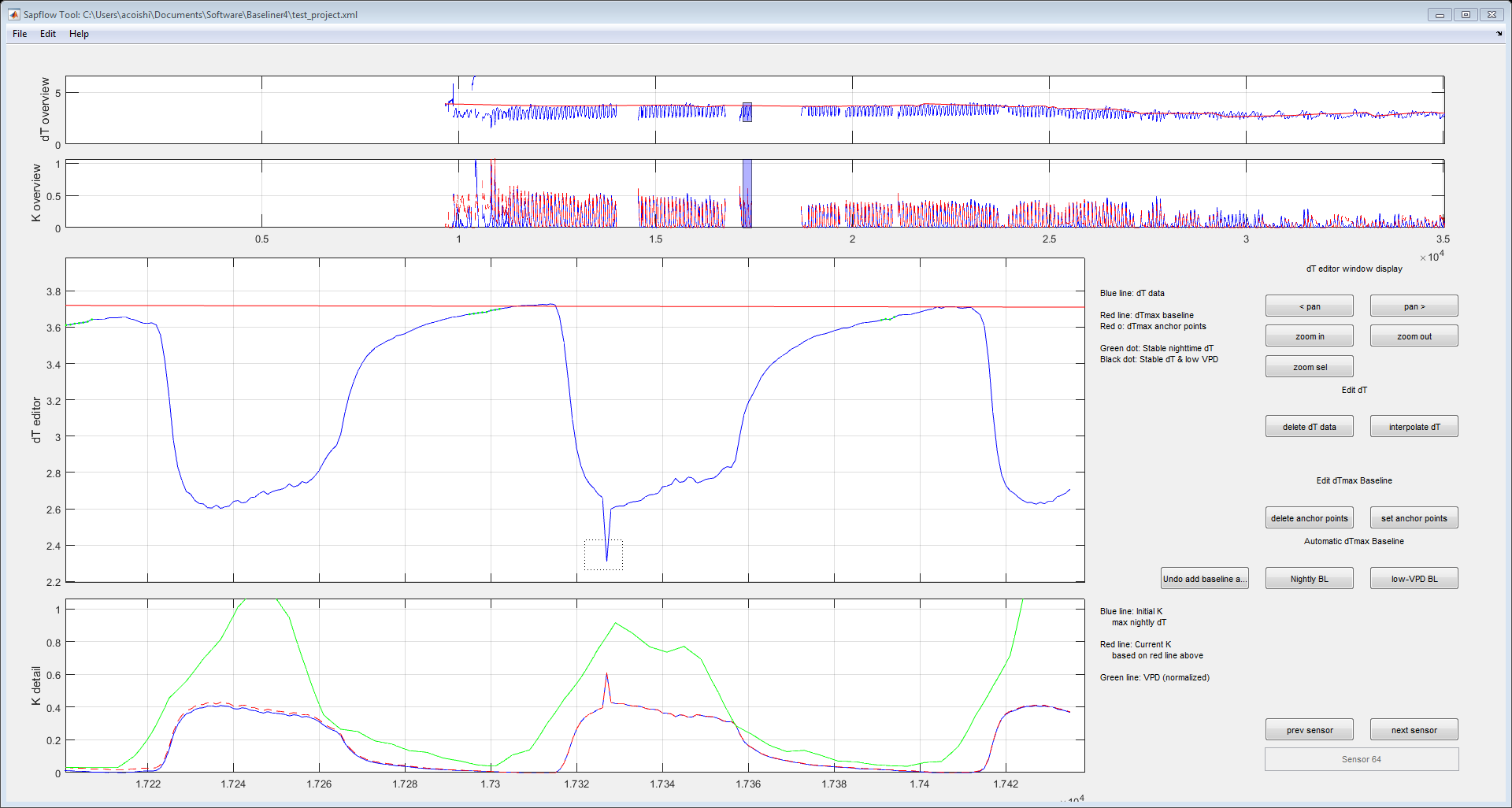
**5.3.6 Interpolating dT Data**

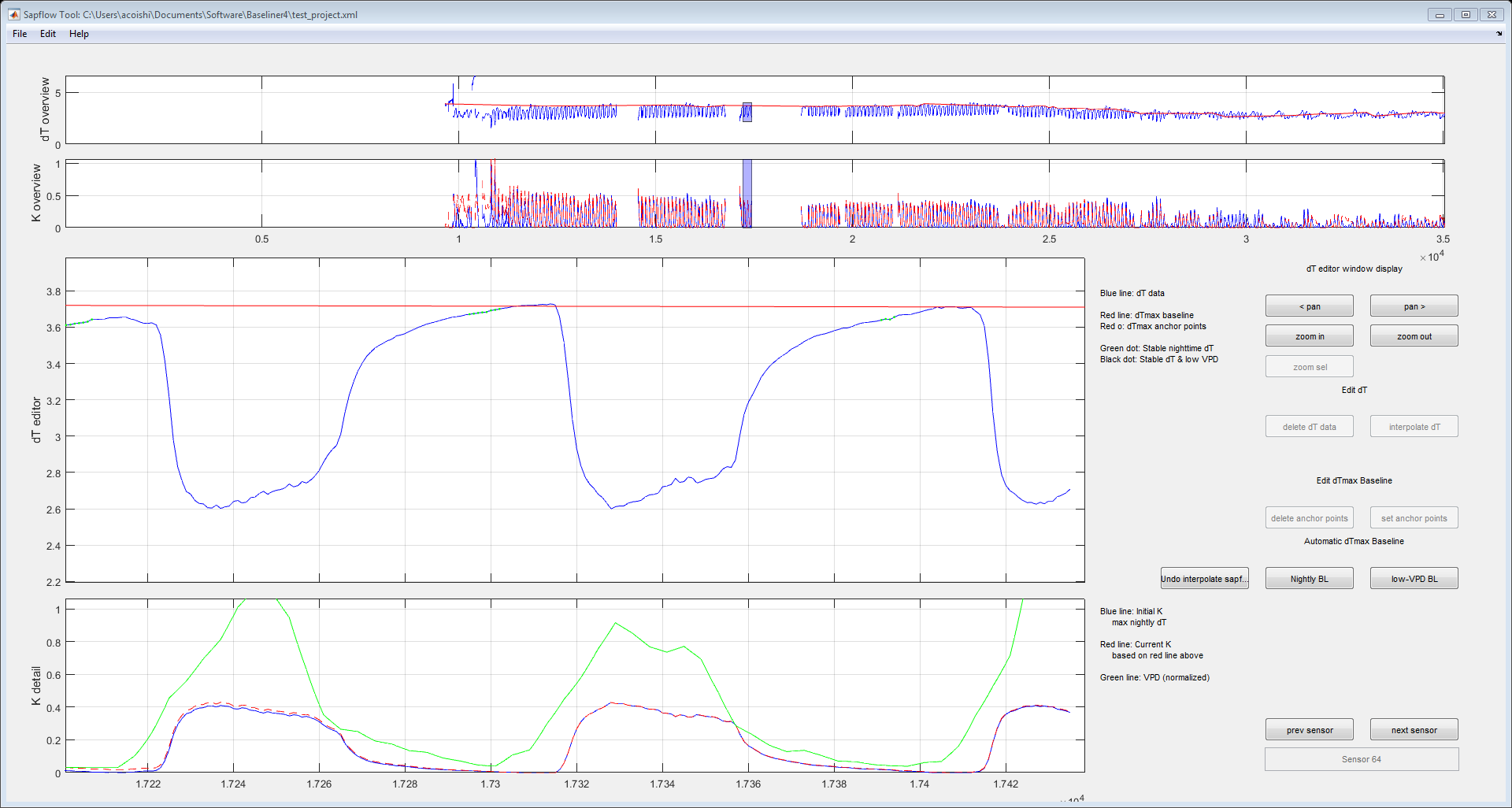
The “Interpolate SF” function can be used to fill small data gaps. You can either drag a box around the area spanning the starting and end points to interpolate between (note that any points between the start and end of the box will be overwritten) or you can right-click within the data gap and the program will automatically draw a box spanning the gap. Once the range of interpolation is selected, click the “Interpolate SF” button and dT data will be filled in. Note that if you click too closely to the blue dT line, the program may try to add that as an anchor point, rather than selecting a box.





The interpolate function can also be used to smooth over data spikes, using the same steps. Note that the selection box only needs to contain the points you want rewritten, so you can position the box in a way that it doesn’t overlap other dT data (i.e. above or below the data).



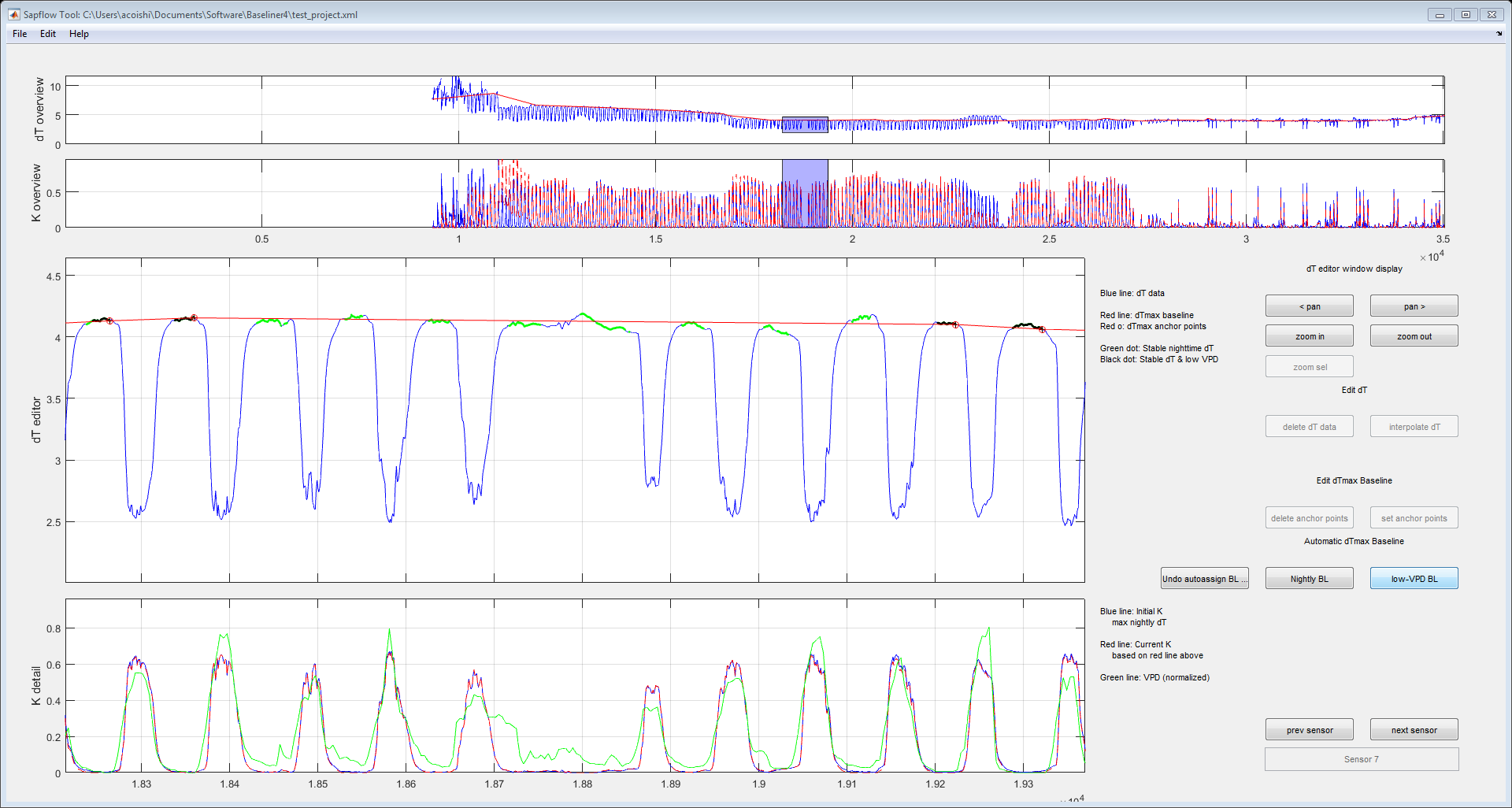


**5.3.8 Undo Operations**

You can undo any operation for the current working sensor using the “Undo” button. Baseliner stores the sequence of operations, so you can sequentially work backwards to undo modifications to dT and manipulation of the dTmax anchor points. After moving to a new sensor, the undo operation is not available for previous sensors.

**5.3.9 Other Considerations**

In many cases, if a baseline anchor is not automatically selected, the interpolated baseline will run very closely through the nightly maximum points. Instances like this are good indicators that (1) the sensor is behaving well and (2) the threshold values for automatically selecting dTmax anchor points are fairly conservative but appropriate, such that the method will only select anchor points that are very likely to correspond with zero flow conditions, skipping over other possible dTmax points.



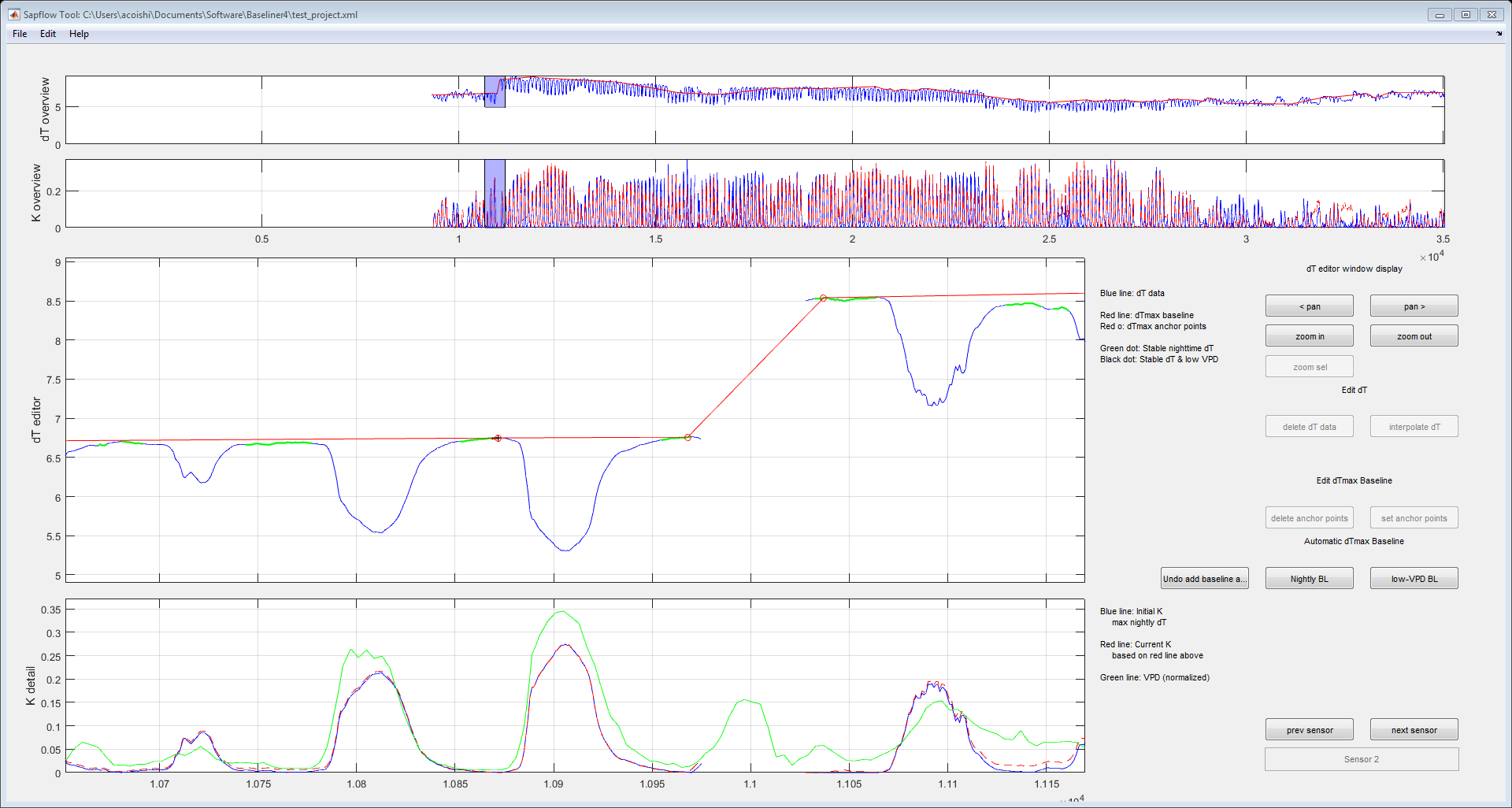
When a modification is made to a probe and/or power supply during the day, dT data are likely to show a sudden step-change.



In these instances, data from the day the step-change occurs are generally not reliable and should be deleted.

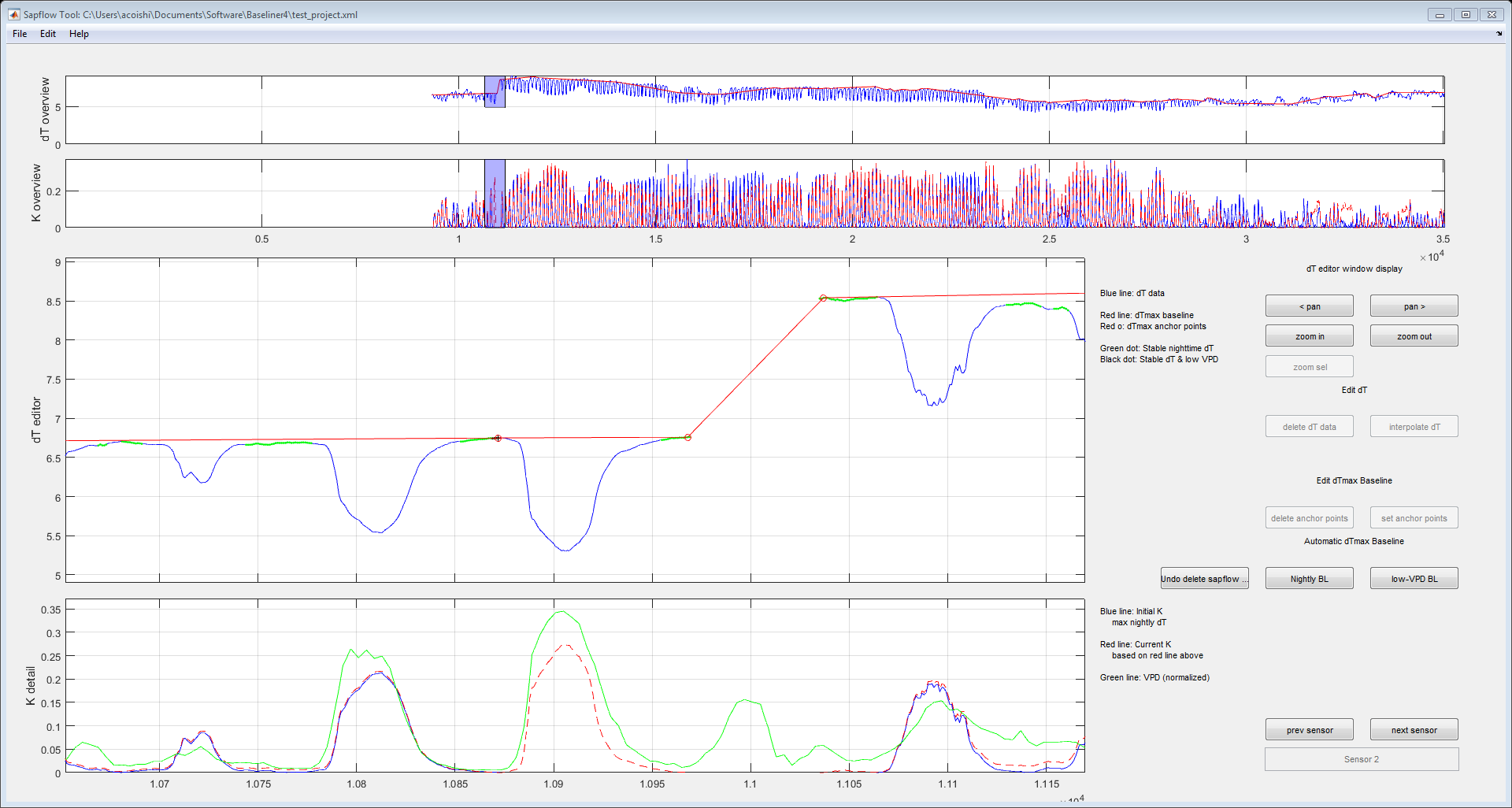


Next new dTmax anchor points must be set, before and after the new data gap.



Finally, since the dTmax baseline changes rapidly the dT data between the anchor points and the gap should be deleted, to prevent exaggerating the calculated K value.

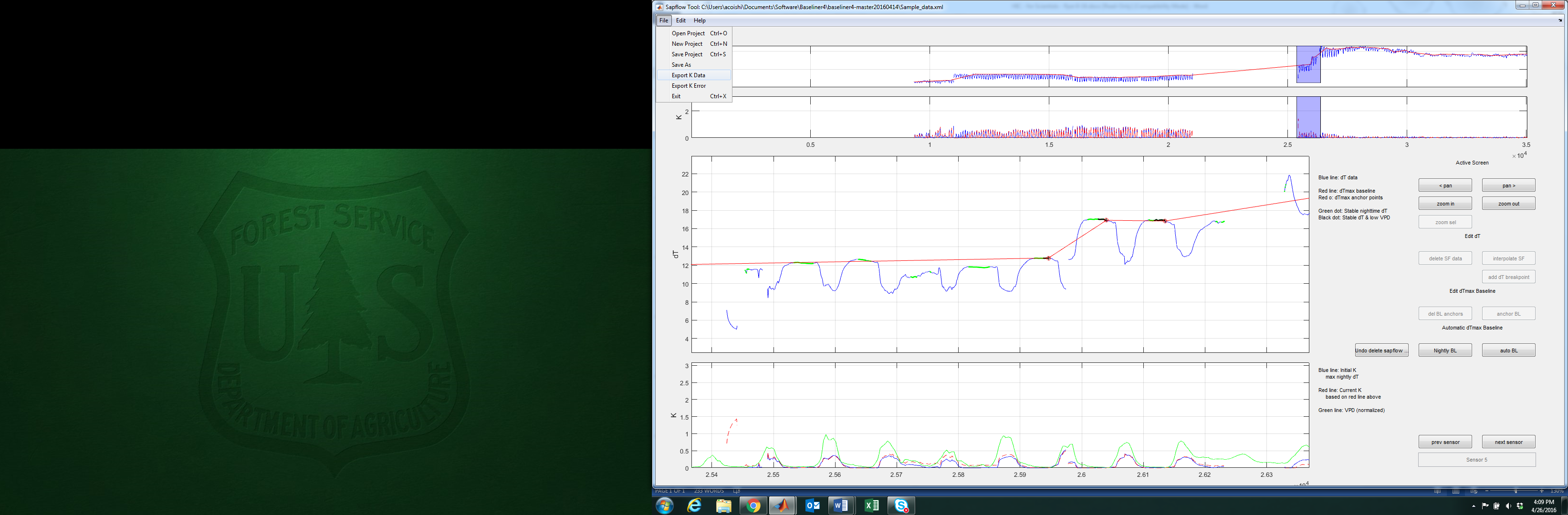




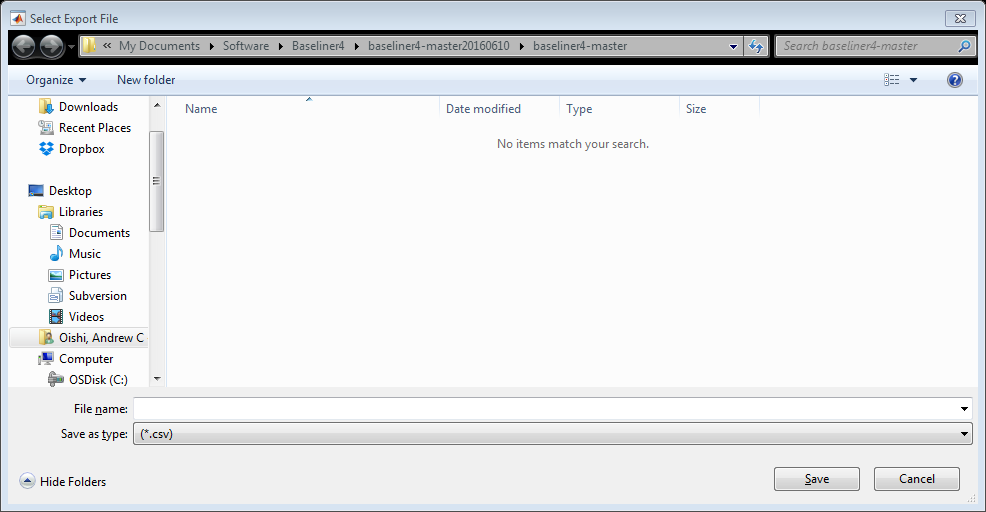
**6. Saving and Exporting data**

When you’ve completed working on a particular sensor, you can click the “next sensor” button and proceed. You can always go back and forth between sensors and continue processing data.

At any time, you can save your work by selecting the “Save Project” option. Frequently saving your work is recommended, to prevent the loss of any data in the event of a computer or software crash.



Once you’ve completed converting data, you can select the “Export K data” option and save your processed data as a .csv file. Column order will be the same as your input file, with dT values replaced by K values. You can export your K data at any point in time (not all sensors need to be completed).



The “Export K error” option will estimate the error in K associated with user error in manually selecting specific dT anchor points. This procedure runs Monte-Carlo simulations of the K estimates based on randomly selected dTmax points, varying around the selected dTmax anchor points within a normal distribution with a standard deviation of 1-hour (independent of time step). The output from this file is the standard deviation of exported K. Columns follow the raw input file.

**References**

Daley MJ and Phillips NG. 2006 Interspecific variation in nighttime transpiration and stomatal conductance in a mixed New England deciduous forest. *Tree Physiology*. 26, 411-419.

Dawson TE, Burgess SSO, Tu KP, Olivera RS, Santiago LS, Fisher JB, Simonin KA, Ambrose AR. 2007. Nighttime transpiration in woody plants from contrasting ecosystems. *Tree Physiology*. 27, 561-575.

Granier A. 1985. Une nouvelle méthode pour la mesure du flux de sève brute dans le tronc des arbres. *Annales des Sciences Forestières*. 42, 193-200

Granier A. 1987. Evaluation of transpiration in a Douglas-fir stand by means of sap flow measurements. *Tree Physiology*. 3, 309-320.

Lu P, Urban L, Zhao P. 2004. Granier’s thermal dissipation probe (TDP) method for measuring sap flow in trees: theory and practice. *Acta Botanica Sinica*. 46, 631-646.

Oishi AC, R Oren, PC Stoy. 2008. Estimating components of forest evapotranspiration: A footprint approach for scaling sap flux measurements. *Agricultural and Forest Meteorology*, 148, 1719-1732.

Oishi AC, Hawthorne D, Oren R. 2016. Baseliner: An open-source, interactive tool for processing sap flux data from thermal dissipation probes. *SoftwareX*. In press.

**Table 1**

Glossary of terms

|  |  |  |
| --- | --- | --- |
| Abbreviation | Description | Units |
| dT | Temperature differential between heated and reference probe | Degrees or mV |
| dTmax | Maximum temperature differential between heated and reference probe when sap flux = 0 | Degrees or mV |
| dTmax anchor point | Discrete points corresponding to recorded dT data where zero-flow sap flux conditions have been identified. The dTmax “baseline” interpolates between these points. | Degrees of mV |
| *F* | Sap flux density | m3 m-2 s-1 |
| *K* | Flow index. Relative difference between dT and dTmax (see Eq 1 and 2) | Unitless |
| PAR | Photosynthetically active radiation |  |
| TDP | Thermal dissipation probe | n/a |
| VPD | Vapor pressure deficit | kPa |
| α | Empirical coefficient relating TDP data to sap flux (see Eq 1) | m3 m-2 s-1 |
| β | Empirical coefficient relating TDP data to sap flux (see Eq 1) | unitless |