A Quantitative Approach for Making Qualitative Assessment Judgments

C. B. Davis and G. A. Roth

2014 IRUG Meeting
Idaho Falls, ID
September 11-12, 2014

www.inl.gov

Idaho National Laboratory

Outline

- Background
  - The traditional, qualitative method for making code assessment judgments is described.
  - A quantitative method for making assessment judgments based on the traditional criteria is proposed.
- Method
  - The quantitative method is described.
- Results
  - Sample results from actual assessment calculations are shown.
- Conclusions

Background

- The Code Scaling, Applicability and Uncertainty (CSAU) process is often used to demonstrate the adequacy of thermal-hydraulic computer codes for simulating accidents in nuclear power plants.
- The CSAU process uses experimental data from integral test facilities to determine code acceptability.
- A standardized and consistent set of criteria has been used in the assessment of codes sponsored by the US Nuclear Regulatory Commission.
- Traditionally, the criteria have been applied subjectively to make qualitative assessment judgments for comparisons between code calculations and measured data.

Background (cont'd)

- The assessment criteria are difficult to apply because they are subjective and vague.
  - Different people could easily come up with different assessment judgments for the same comparison with data.
  - Different features on the same plot could lead to different judgments based on conflicting criteria.
  - No guidance is given to resolve conflicts.
- The assessment judgments of excellent, reasonable, minimal, and insufficient are defined on subsequent slides.

*Excellent agreement**

- Applies when the code exhibits no deficiencies in modeling a given behavior. Major and minor phenomena and trends are correctly predicted. The calculated results are judged to agree closely with the data. **The calculation will, with few exceptions, lie within the uncertainty bands of the data.** The code may be used with confidence in similar applications. (The term “major phenomena” refers to the phenomena that influence key parameters such as fuel rod cladding temperature, pressure, differential pressure, mass flow rate, and mass distribution. Predicting major trends means that the prediction shows the significant features of the data. Significant features include the magnitude of a given parameter through the transient, slopes, and inflection points that mark significant changes in the parameter.)

---

* R. R. Schultz, International Code Assessment and Applications Program: Summary of Code Assessment Studies Concerning RELAP5/MOD2, RELAP5/MOD3, and TRAC-B, Idaho National Engineering Laboratory, NUREG/IA-0128, EGG-EAST-8719, December 1993.

Reasonable agreement

- Applies when the code exhibits minor deficiencies. Overall, the code provides an acceptable prediction. All major trends and phenomena are correctly predicted. Differences between calculation and data are greater than deemed necessary for excellent agreement. **The calculation will frequently lie outside but near the uncertainty bands of the data.** However, the correct conclusions about trends and phenomena would be reached if the code were used in similar applications. The code models and/or facility model noding should be reviewed to see if improvements can be made.

Minimal agreement

- Applies when the code exhibits significant deficiencies. Overall, the code provides a prediction that is only conditionally acceptable. Some major trends or phenomena are not predicted correctly, and some calculated values lie considerably outside the uncertainty bands of the data. Incorrect conclusions about trends and phenomena may be reached if the code were used in similar applications, and an appropriate warning needs to be issued to users. Selected code models and/or facility model noding need to be reviewed, modified, and assessed before the code can be used with confidence in similar applications.

Insufficient agreement

- Applies when the code exhibits major deficiencies. The code provides an unacceptable prediction of the test. Major trends are not predicted correctly. Most calculated values lie outside the uncertainty bands of the data. Incorrect conclusions about trends and phenomena are probable if the code is used in similar applications, and an appropriate warning needs to be issued to users. Selected code models and/or facility model noding need to be reviewed, modified, and assessed before the code can be used with confidence in similar applications.

Background (cont'd)

- Assessment judgments of “excellent” or “reasonable” are considered to indicate acceptable code performance.
- There are no acceptance requirements for the assessment; it is not required that a certain number of parameters have acceptable judgments.

Background (cont'd)

– The assessment judgments depend heavily on the measurement uncertainty.
– Unfortunately, reported measurement uncertainty is actually just an estimate of the uncertainty, and often varies significantly between experiments and facilities.
– Independent measurements can often be used to show that the performance of an actual instrument is either much better or much worse than its reported uncertainty.
– Therefore, engineering judgment needs to be used in the interpretation of reported uncertainties.

Background (cont'd)

- INL assessed RELAP5-3D using data from various integral experiments to support the licensing of the mPower reactor.
- Eight integral experiments from the Loss-of-Fluid Test (LOFT) and Semiscale facilities were assessed.
  - These experiments simulated small-break loss-of-coolant accidents and anticipated operational occurrences.
  - Approximately 150 comparison plots were reported.
  - The use of the traditional, qualitative approach to assessment judgments would be time consuming.
- Therefore, a more quantitative approach was used to help make the required assessment judgments.

A quantitative approach was used to help make the qualitative assessment judgments

- The quantitative approach utilized the estimated uncertainty and the difference between the calculated and measured results.
- The absolute error at each minor edit time point in the RELAP5-3D calculation was calculated based on the absolute value of the difference between the calculated and measured parameters.
- A normalized error was then obtained by dividing the absolute error by the estimated uncertainty.
    - The uncertainty was taken to be 2 standard deviations.

A quantitative approach was used to help make the qualitative assessment judgments (cont'd)

- The fraction of time points when the normalized error was less than a given factor was then determined.
  - To achieve an excellent judgment, at least 90% of the normalized errors must be less than one, which corresponds to the calculated parameter being within the data uncertainty 90% of the time.
  - To achieve a reasonable judgment, at least 70% of the points must lie within three times the uncertainty, which corresponds to the calculated parameter being within three times the uncertainty 70% of the time.
    - The factor of three quantifies the meaning of "near" and the 70% quantifies "frequently", which are used in the definition of "reasonable".

A quantitative approach was used to help make the qualitative assessment judgments (cont'd)

- The model parameters (90%, 70%, and 3 times the uncertainty) are intuitively reasonable and produce assessment judgments that are generally consistent with subjective evaluations.
- Plots that did not meet the excellent or reasonable criteria were assigned judgments of minimal.
- The quantitative method provided a starting point for making the assessment judgments, but engineering judgment was still required.
- The results from the quantitative model were visually checked and judged to be reasonable.
  - In theory, the quantitative judgments could be overridden by engineering judgment, but no such adjustments were performed for the LOFT and Semiscale assessments.
- The use of the quantitative model allowed the assessment judgments to be made easily and quickly.

Quantitative method

- An AptPlot script was used to generate the assessment figures for each experiment.
  - The figures contained estimated uncertainty values, which were taken from, or derived from, the experiment data reports.
  - The AptPlot script generated two files for each figure, a jpeg file that was included in the report and an apf file (in ascii) that contained all the information necessary to make the figure, including the calculated and measured data and the estimated uncertainty.

Quantitative method (cont'd)

- Two simple Fortran programs were written to implement the quantitative approach for making the qualitative assessment judgments.
- The first program read the .apf file that was created by an AptPlot script and wrote the calculated and measured results in two files that contain X-Y pairs.
- The estimated uncertainty was read from the .apf file and passed to the second program by a Linux script.
- The second program read the two X-Y files and determined the time channel from the X values of the RELAP5-3D calculation.
- The Y values from the experiment were then interpolated to determine appropriate values at the time points from the RELAP5-3D calculation.

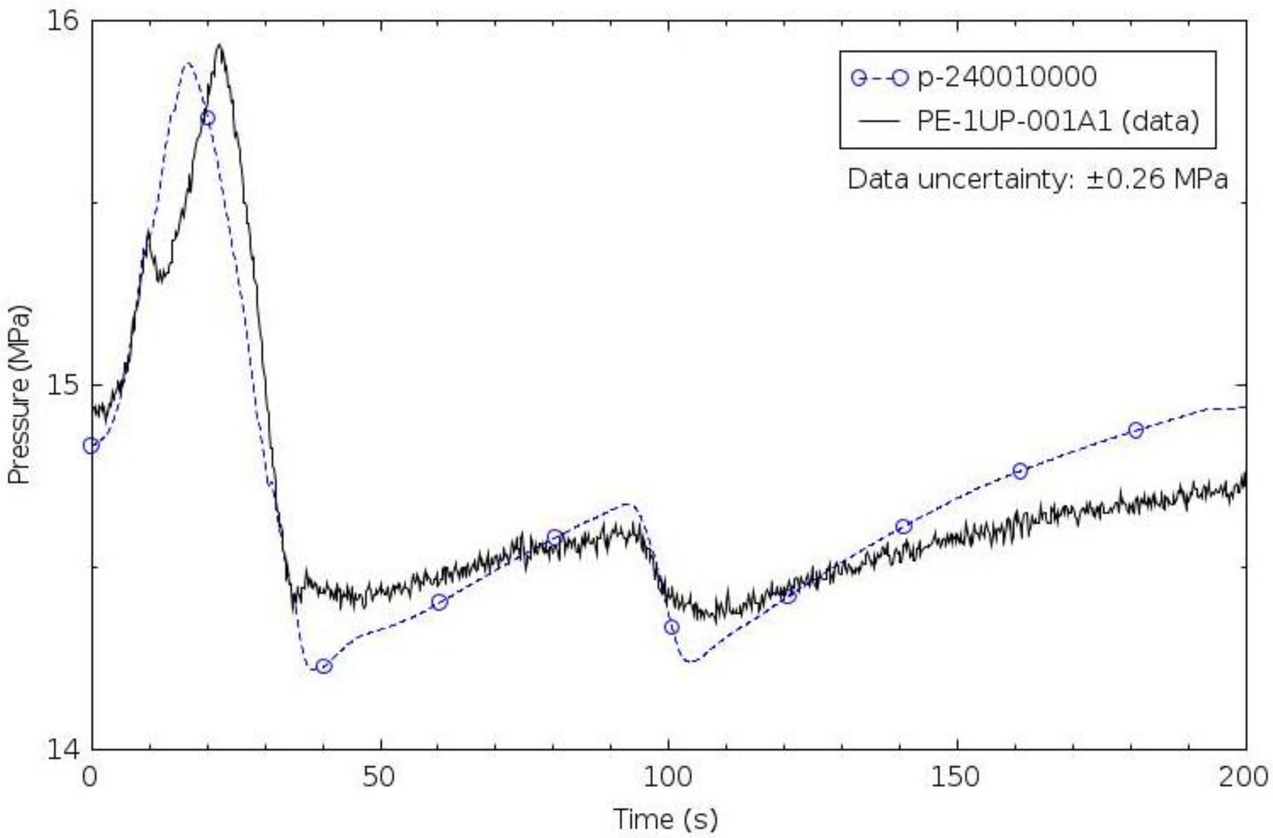
Quantitative method (cont'd)

- The second program then compared the Y values from the calculated and measured sets and determined the fraction of time points that are within one times and three times the estimated uncertainty.
- Only times that contained both calculated and measured data were considered.
  - For example, measured data after the end of the calculation were ignored.

Results from some of the INL assessment calculations performed for mPower follow

- LOFT
  - Experiment L3-1 simulated a 4-inch cold leg break.
  - Experiment L6-1 simulated a loss of steam load.
  - Experiment L6-2 simulated a loss of forced coolant flow in the primary system.
- Semiscale Test S-07-10 simulated a 10% cold leg break.
- Example figures are presented that show different judgments.
- The assessment judgments correspond to the time scale shown in the figure, not the duration of the calculation.
- Each figure shows:
  - F_{unc}, the fraction of calculated points within the estimated uncertainty.
  - F_{3unc}, the fraction of calculated points within 3 times the estimated uncertainty.

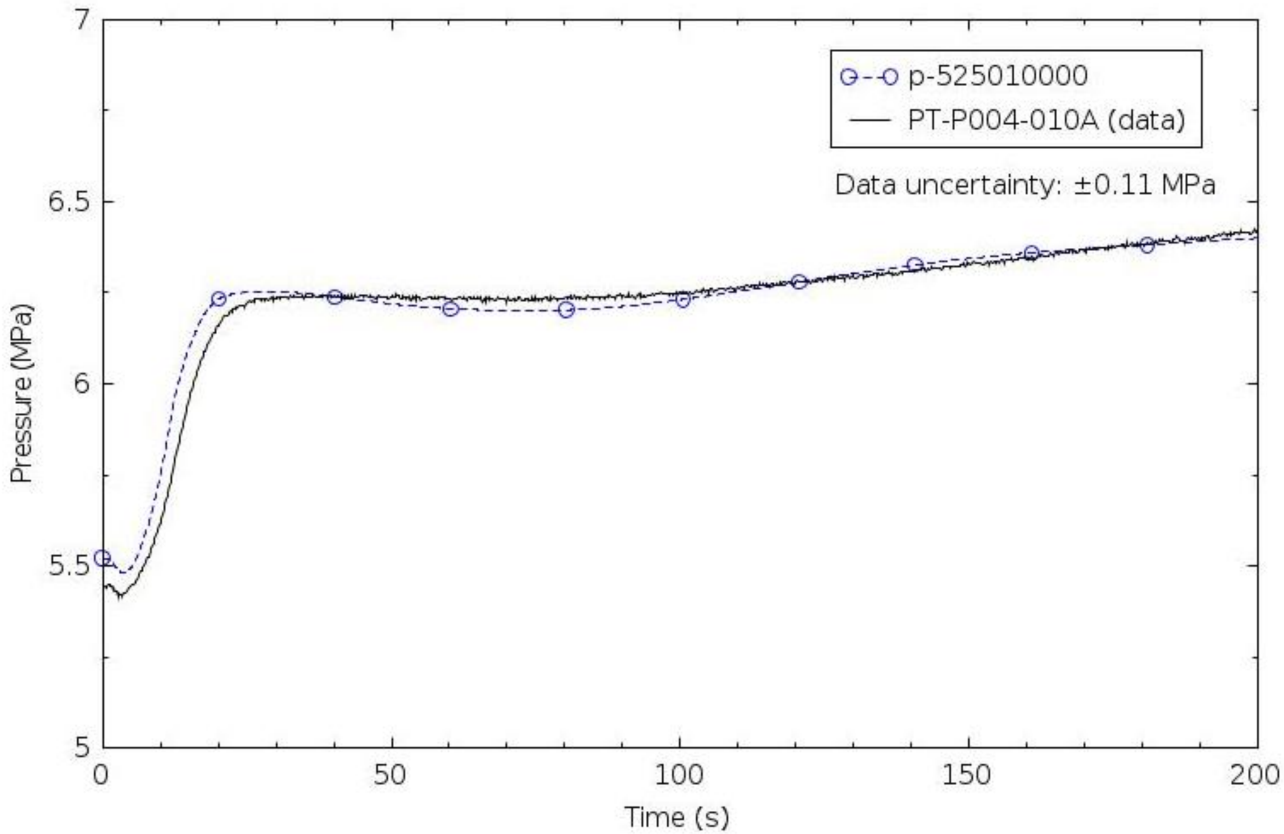
LOFT L6-1 PCS Pressure: Excellent



$F_{unc} = 0.93$

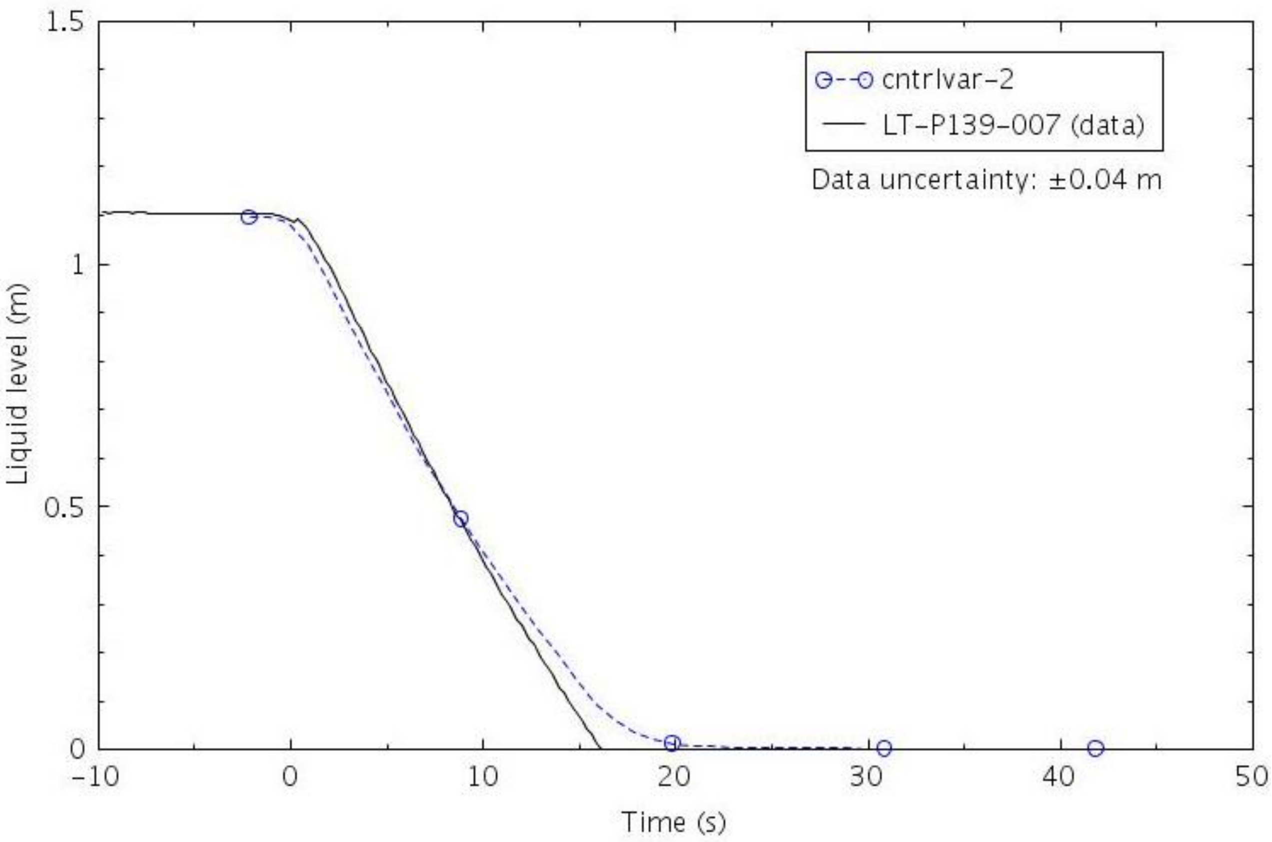
$F_{3unc} = 1.00$

LOFT L6-2 SG Pressure: Excellent



$F_{unc} = 0.96$
$F_{3unc} = 1.00$

LOFT L3-1 Pressurizer Level: Reasonable

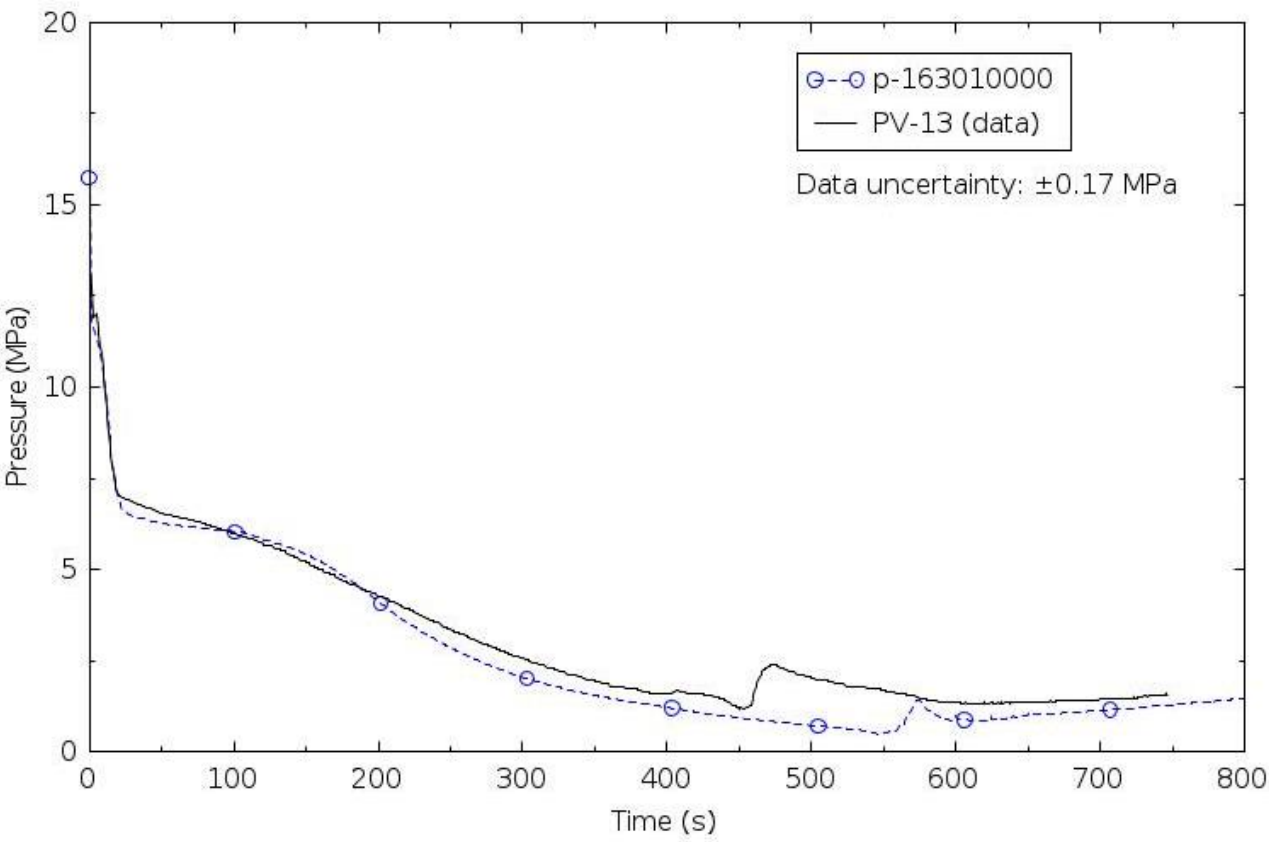


$F_{unc} = 0.89$

$F_{3unc} = 1.00$

The time scale should be reasonable based on the phenomena of interest.

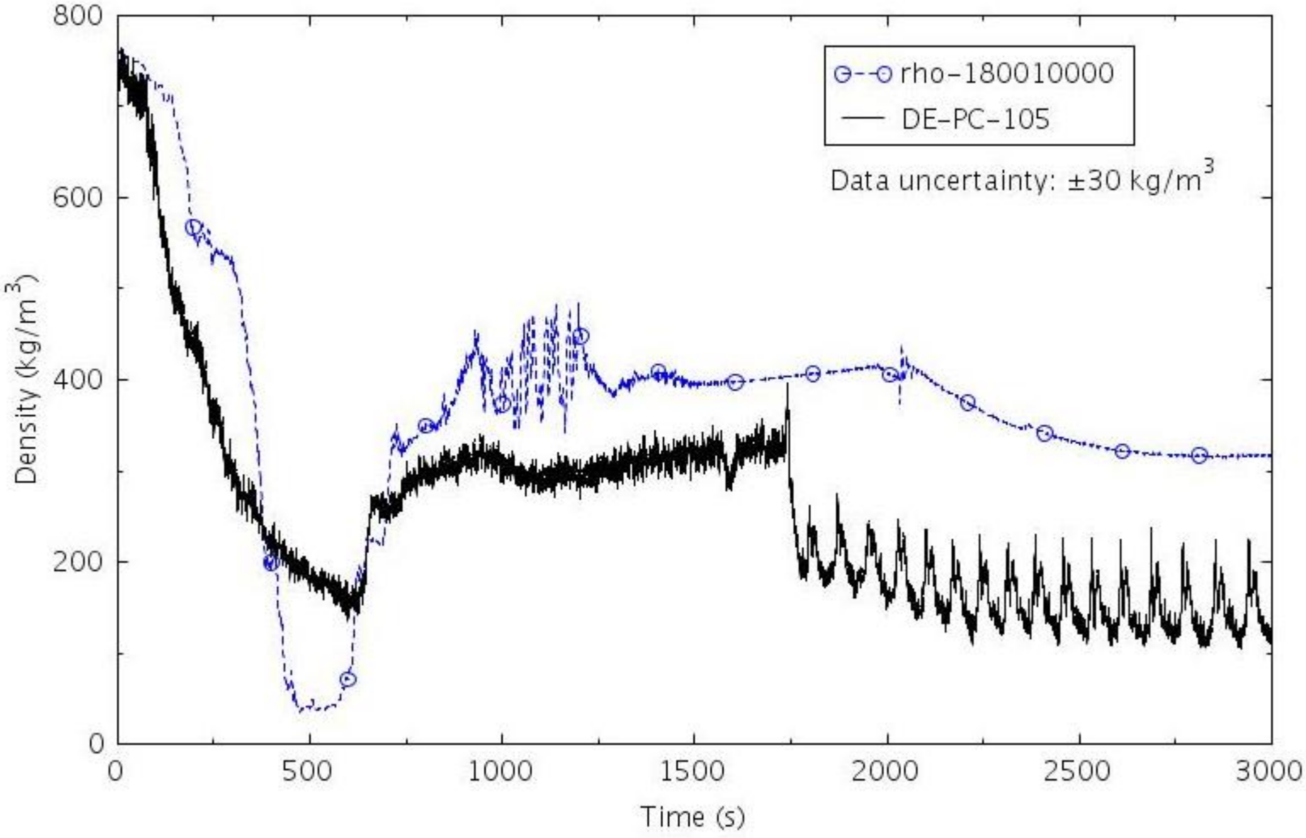
Semiscale S-07-10 PCS Pressure: Reasonable



$F_{unc} = 0.15$

$F_{3unc} = 0.79$

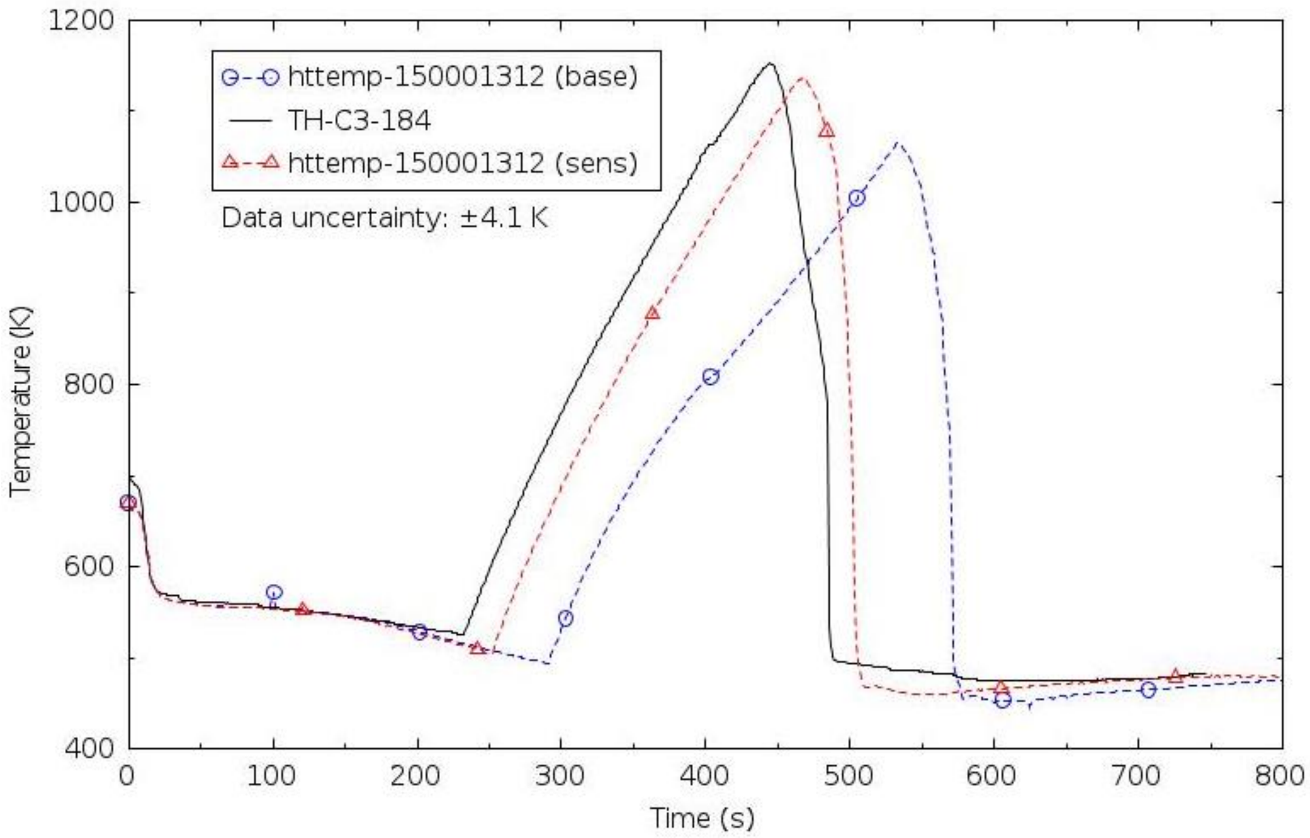
LOFT L3-1 Intact Loop Cold Leg Density: Minimal



$F_{unc} = 0.04$

$F_{3unc} = 0.29$

Semiscale S-07-10 Cladding Temperature: Minimal



Base:

$F_{unc} = 0.23$

$F_{3unc} = 0.33$

Sensitivity:

$F_{unc} = 0.34$

$F_{3unc} = 0.52$

A minor timing offset combined with a small uncertainty can significantly lower qualitative judgments.

The variability between similar measurements during the dryout is much greater than the estimated uncertainty.

Additional Points to Consider

- Time scales should be reasonable based on the physics of the phenomena.
- A uniform time scale should be used to avoid biasing the results towards periods of high plot point density.
- Offsets in time can cause lower than expected rankings.
- The quantitative approach does not specifically evaluate differences in slopes.
  - Differences in slope will affect assessment judgments if they persist long enough.

Conclusions

- The definitions of assessment judgments used in the CSAU process are qualitative, subjective, difficult to interpret, and can result in different rankings by different people.
- A more quantitative approach was developed that is based on the CSAU definitions.
- The quantitative approach utilized the estimated uncertainty to provide qualitative judgments.
- The results of the quantitative approach are generally consistent with engineering judgment.
- The CSAU definitions and results depend heavily of the estimated measurement uncertainty.
    - Poor uncertainty estimates affect assessment judgments.