May-2024

Another Look at the Yoshida Correlation for Predicting Impact Sensitivity

E. Margelefsky

Process R&D (Environmental & Process Safety Eng'g Group)

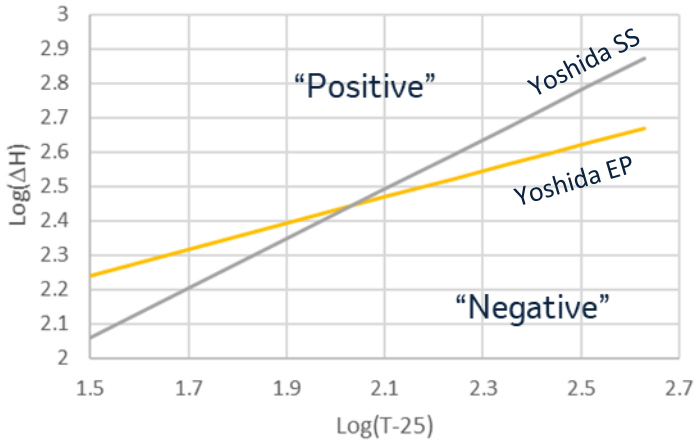
Merck & Co., Inc., Rahway, NJ, USA

Background

Some of the compounds we handle in R&D and/or manufacturing are potentially explosive and/or shock sensitive.

- We can't do full suite of explosivity tests on every compound
- Yoshida (1987) generated a correlation to predict **explosion propagation (EP)** and **shock-sensitivity (SS)** based on DSC data alone
- According to IQ Survey, 7 of 10 pharma companies surveyed reported using the Yoshida correlations.
- Incorporated into ASTM E1231 "Standard Practice for Calculation of Hazard Potential Figures-of-Merit for Thermally Unstable Materials" and referenced in CSB reports. **Widely accepted/used.**

Org. Process Res. Dev. 2020, 24, 11, 2529−2548



$$EP = \log(\Delta H) - 0.38 \log(T\text{-}25) - 1.67$$

$$SS = \log(\Delta H) - 0.72 \log(T\text{-}25) - 0.98$$

Yoshida, Tadao, et al. "Prediction of fire and explosion hazards of reactive chemicals (Part 1). Estimation of explosive properties of self-reactive chemicals from SC-DSC data." *Kogyo Kayaku;(Japan)* 48.5 (1987).

Background

As a frequent user of this correlation, I found myself wondering…

- How much data was used in the construction of these correlations? How robust are they?
- Has anyone independently confirmed/re-assessed the original correlation?
- What exactly did Yoshida mean by “shock-sensitive” and “explosion propagating”?
- Are the DSC parameters really the best predictor variables?
- **Can we do better** in the age of “big data” machine learning?

Background

There are various examples in the literature of people publishing DSC data and Yoshida predictions with no confirmatory testing, or applying the technique to systems beyond the initial scope of pure substances.

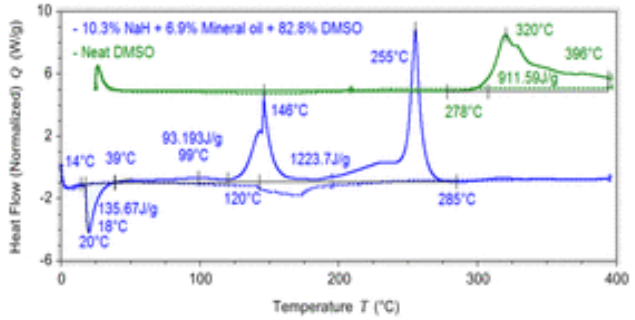
- Overstatement of the results:

> determination of the exothermic peak temperatures for D1, D2, and D3 as 145.3 °C, 145.2 °C, and 148.2 °C, respectively, and the corresponding onset temperatures as 115.8 °C, 116.7 °C, and 118 °C, respectively, could be achieved from the differential scanning calorimetry (DSC) curves. The similar values of the exothermic peaks between D1 and D2 may be attributed to their comparable molecular weights, energy gaps, and enthalpy values. Additionally, they may share similar thermal decomposition reaction mechanisms, such as the decomposition of diazirine moieties. Furthermore, the application of the Yoshida correlation and the DSC confirmed that none of the three dyes were likely to be explosive (as illustrated in Table 1), at which point their synthesis processes could be safely scaled up for large-scale application.

https://doi.org/10.1016/j.dyepig.2023.111784

- Yoshida correlation applied to a reactive mixture of NaH in DMSO:

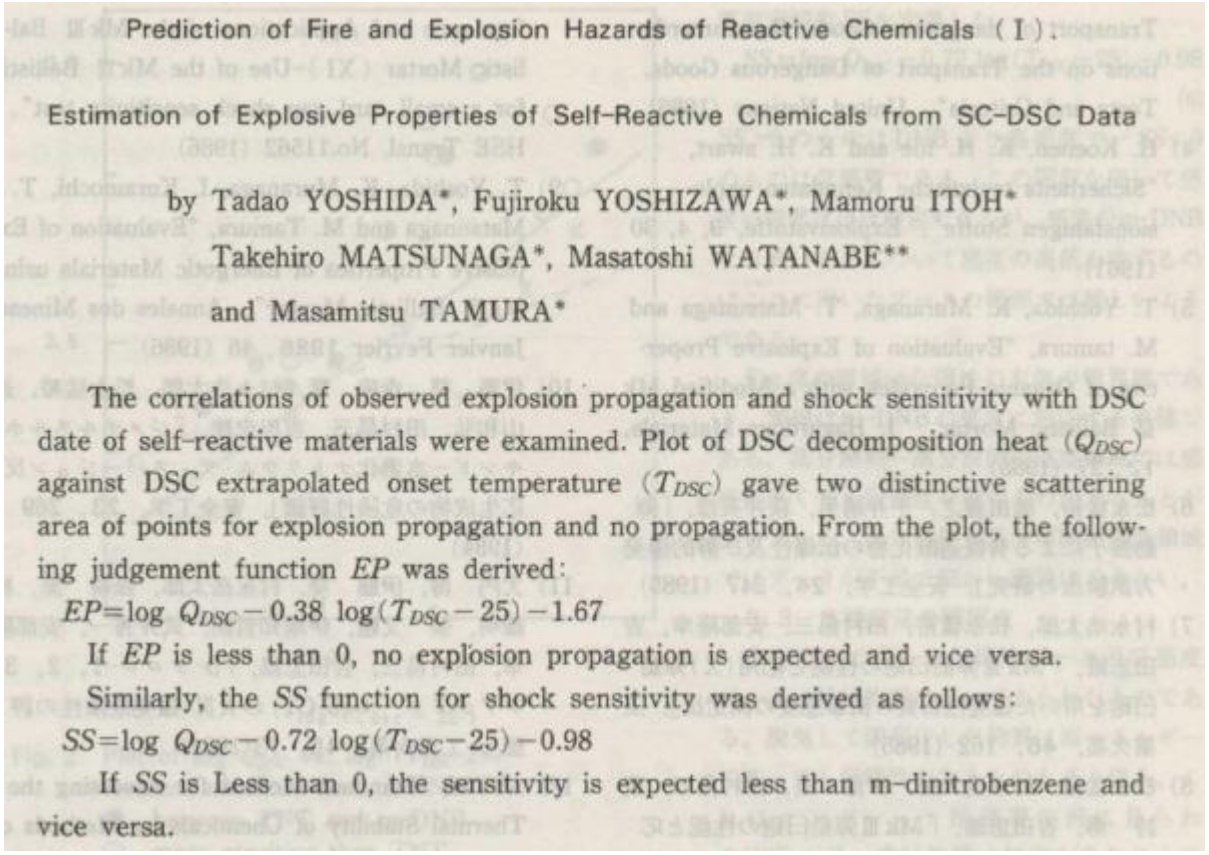
> because of the ARC cell rupture, DSC evaluation of a mixture containing 17.2% NaH in mineral oil (60 wt %) and 82.8% DMSO was performed to better understand the total heat output of this thermal decomposition (Figure 7, left). Initially an endothermic event (14 to 39 °C) was observed, which was immediately followed by a minor exothermic event (39 to 120 °C) with a total energy release of −93 J/g. Starting at 120 °C, two consecutive major exothermic events with a combined heat of −1224 J/g were detected. The sharp narrow peaks in the DSC profile are characteristic of autocatalytic reactions, which represent much more significant hazards of thermal decomposition. These two combined events were deemed to be explosive according to the Yoshida correlation. (53) It is worth noting that



https://pubs.acs.org/doi/10.1021/acs.oprd.0c00159

Original Correlations

- The paper is in Japanese (other than the abstract) and is not available digitally, but a scanned version was obtained through Article Galaxy



Prediction of Fire and Explosion Hazards of Reactive Chemicals (I).

Estimation of Explosive Properties of Self-Reactive Chemicals from SC-DSC Data

by Tadao YOSHIDA*, Fujiroku YOSHIZAWA*, Mamoru ITOH*
Takehiro MATSUNAGA*, Masatoshi WATANABE**
and Masamitsu TAMURA*

The correlations of observed explosion propagation and shock sensitivity with DSC date of self-reactive materials were examined. Plot of DSC decomposition heat (Q_{DSC}) against DSC extrapolated onset temperature (T_{DSC}) gave two distinctive scattering area of points for explosion propagation and no propagation. From the plot, the following judgement function EP was derived:

$EP = \log Q_{DSC} - 0.38 \log(T_{DSC} - 25) - 1.67$

If EP is less than 0, no explosion propagation is expected and vice versa.

Similarly, the SS function for shock sensitivity was derived as follows:

$SS = \log Q_{DSC} - 0.72 \log(T_{DSC} - 25) - 0.98$

If SS is Less than 0, the sensitivity is expected less than m-dinitrobenzene and vice versa.

Yoshida, Tadao, et al. "Prediction of fire and explosion hazards of reactive chemicals (Part 1). Estimation of explosive properties of self-reactive chemicals from SC-DSC data." *Kogyo Kayaku;(Japan)* 48.5 (1987).

Original Correlations

Yoshida's original data set is sparse, and has few non-explosives

EP:

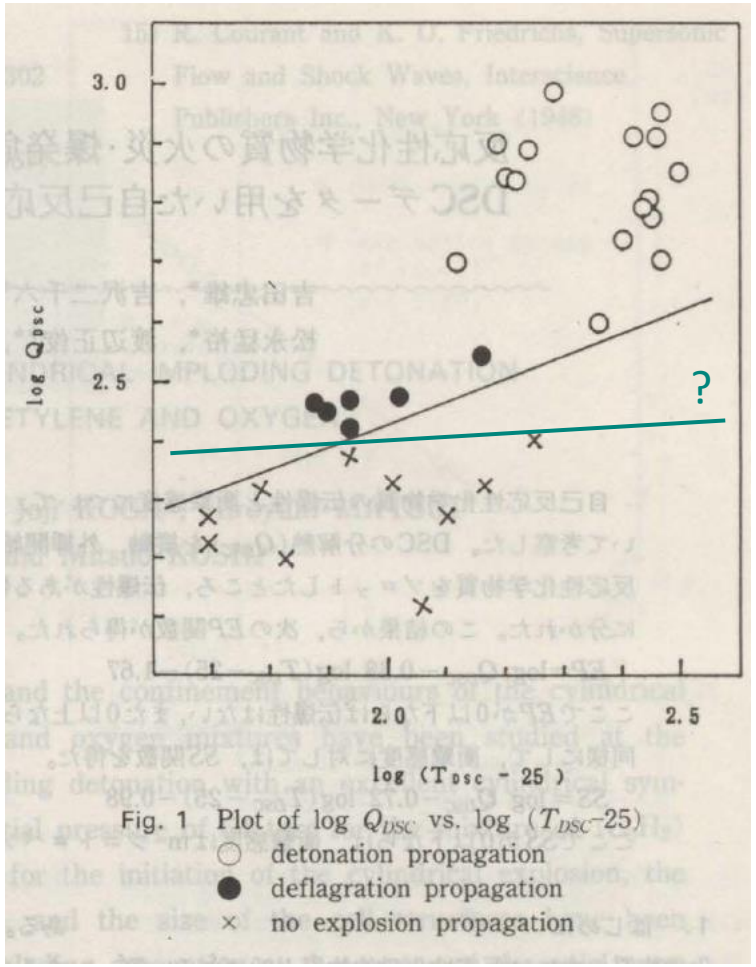


Fig. 1 Plot of log Q_{DSC} vs. log (T_{DSC}–25)
○ detonation propagation
● deflagration propagation
× no explosion propagation

10 non-EP
23 EP

SS:

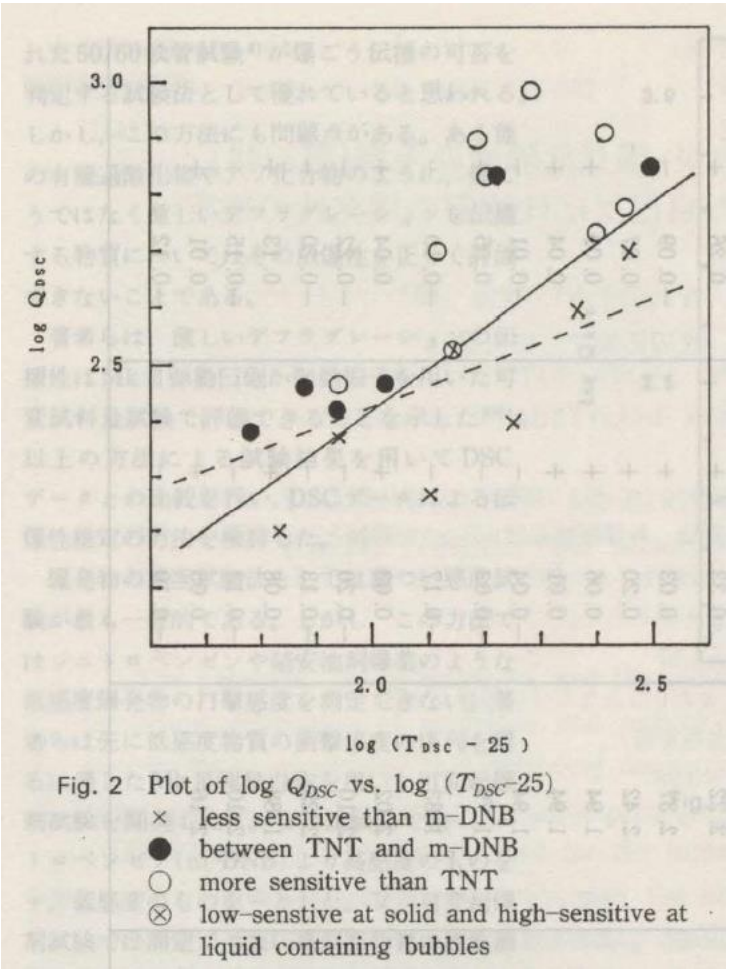


Fig. 2 Plot of log Q_{DSC} vs. log (T_{DSC}–25)
× less sensitive than m–DNB
● between TNT and m–DNB
○ more sensitive than TNT
⊗ low–senstive at solid and high–sensitive at liquid containing bubbles

6 non-SS
15 SS

Original Correlations

Yoshida's original data set is sparse (23 data points), many of which are high explosives

2, 4, 6-Trinitrotoluene (TNT)
2, 4, 6-Trinitrophenol (picric acid)
2, 4, 6-Trinitro-m-xylene
m-Dinitrohenzene (m-DNB)
2, 4-Dinitrotoluene (2, 4-DNT)
2, 4, 6-Diazodinitrophenol (DDNP)
Trimethylenetrinitramine (RDX)
Tetramethylenetetranitramine (HMX)
Dinitrosopentamethylene-tetramine (DPT)
Ammonium nitrate
$(NH_4)_2SO_4 \cdot 2NH_4NO_3$ → Mixture of 2 chemical species
Benzoyl peroxide (BPO)
90% BPO + 10% water } Mixtures of 2 chemical species
80% BPO + 10% water }
Lauroyl peroxide → Presumably a typo
Monoperoxyphthalic acid magnesium hexahydrate
t-Butyl perbenzoate
Dicumyl peroxide
Azodicarbonamide
α, α'-Azobisisobutylonitril (AIBN)
p-Toluenesulfonyl hydrazide
5-Chlorothiadiazole (5-CT)
Sodium dimethylsulfinate → DSC showed 2 distinct exotherms;
Yoshida developed his own averaging technique just for this one data point

SS:

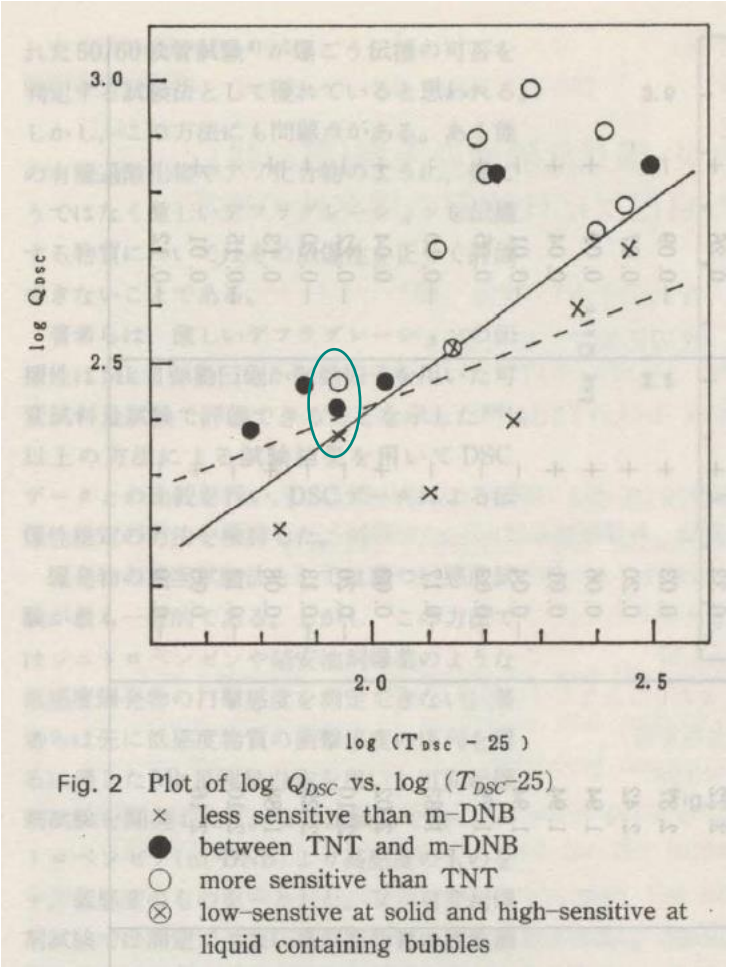


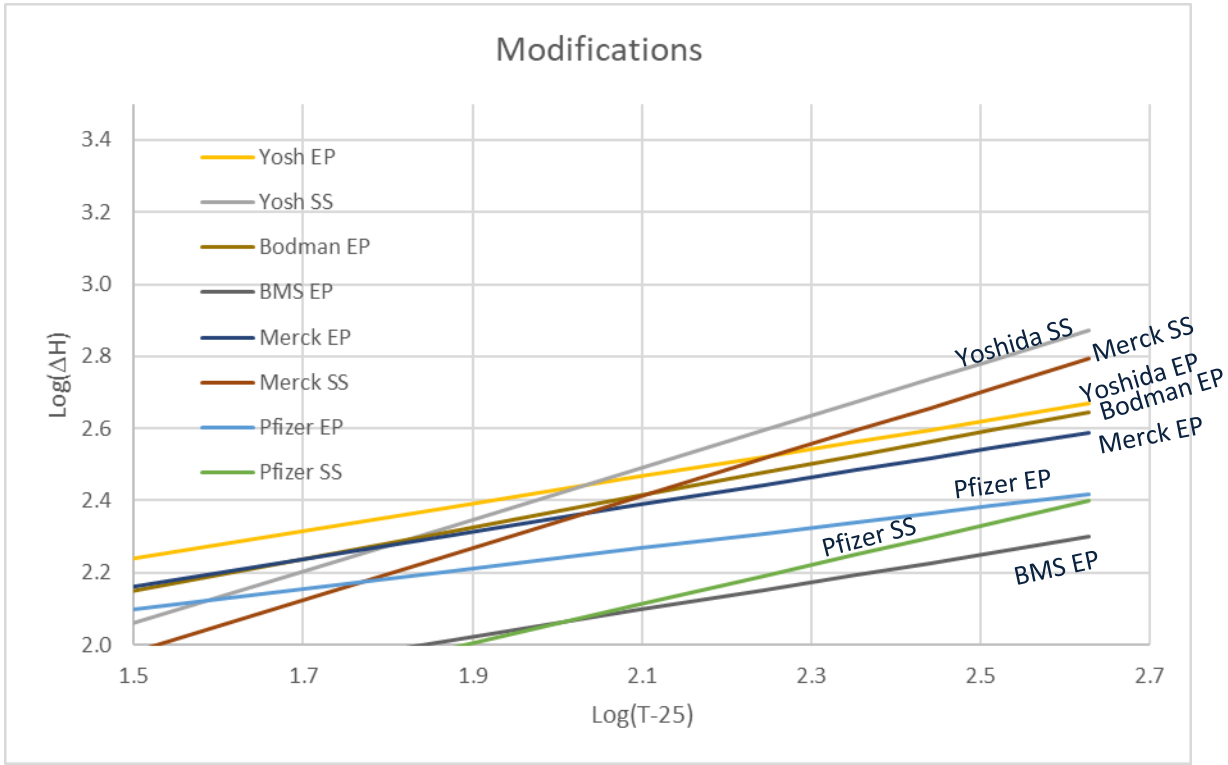
Fig. 2 Plot of log Q_{DSC} vs. log $(T_{DSC}-25)$
× less sensitive than m-DNB
● between TNT and m-DNB
○ more sensitive than TNT
⊗ low-senstive at solid and high-sensitive at liquid containing bubbles

Other Modifications

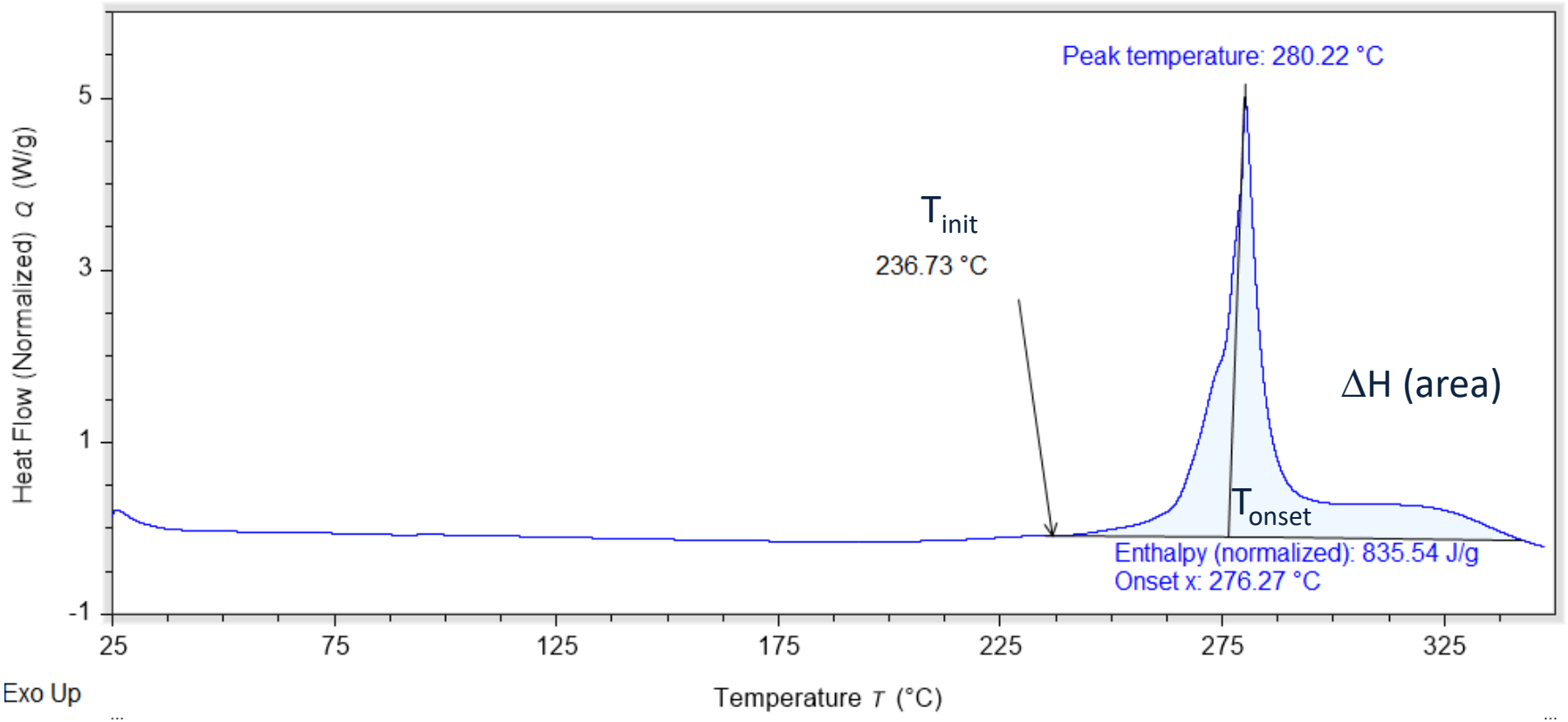
Various companies have developed “modifications” of original Yoshida correlations

- Same approach using log(T-25) and log(ΔH) but shift it to be more conservative; generate fewer false-negatives
- Each one is generated by modifying slope or intercept of the original lines, trying to capture all the positives, largely ignoring negatives.
- Bodman developed a new correlation for EP using a new dataset (n=22)
- Most pharmas also use T_{init} rather than T_{onset}

- At Merck, if a material is flagged as “Yoshida positive” we would usually follow up with a drop-weight test (BAM Fallhammer) which directly tests for shock-sensitivity
  - Explosivity testing is less common and must be outsourced. Several attempts to “validate” EP correlation have been attempted (largely inconclusive)
  - SS data is easier to generate and find, and is often more useful for pharma R&D purposes
  - **Our current focus is on SS rather than EP**



DSC – initiation vs. onset temp



- T_{init} is where exotherm first deviates from baseline
- T_{onset} is where steepest tangent line intersects baseline (used by Yoshida)

A second Yoshida EP Correlation?

The original paper was “Part 1” of a series. In Part 4 Yoshida focused on organic peroxides and mixtures:

“Prediction of fire and explosion hazards of reactive chemicals (Part 4). Estimation of explosive property of organic peroxides from SC-DSC”

The Correlation of observed explosion propagation data with DSC data of self-reactive materials were examined. Plot of logarithum of DSC decomposition heat (Q_{DSC}) against that of DSC extraporated onset temperature (T_{DSC}) gave two distinctive scattering area of points regarding the ability to propagate an explosion. From the plot, the following judgement function EP was derived:

$$EP = \log Q_{DSC} - 0.57 \log (T_{DSC} - 25) - 1.34$$

5 new data points → New “Correlation”!

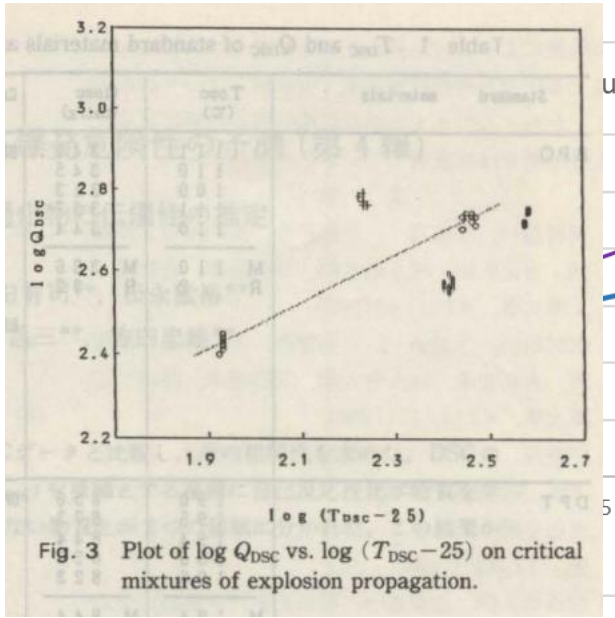
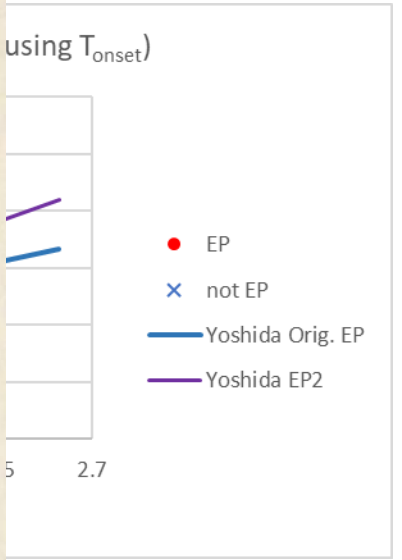


Fig. 3 Plot of log Q_{DSC} vs. log (T_{DSC}−25) on critical mixtures of explosion propagation.



- The fact that Yoshida created a second correlation rather than using or updating the original EP correlation seems to undermine the value of the initial correlation

Watanabe, Wada, Matsunaga, Itoh, Tamura, and Yoshida. Prediction of fire and explosion hazards of reactive chemicals (Part 4). Estimation of explosive property of organic peroxides from SC-DSC. Kogyo Kayaku, Vol 50 No.2 pp100-105.

Can we do better today?

With Ben Dobson (AZ), I compiled a dataset of compounds for which both DSC and Fallhammer data exists, using both published and unpublished data

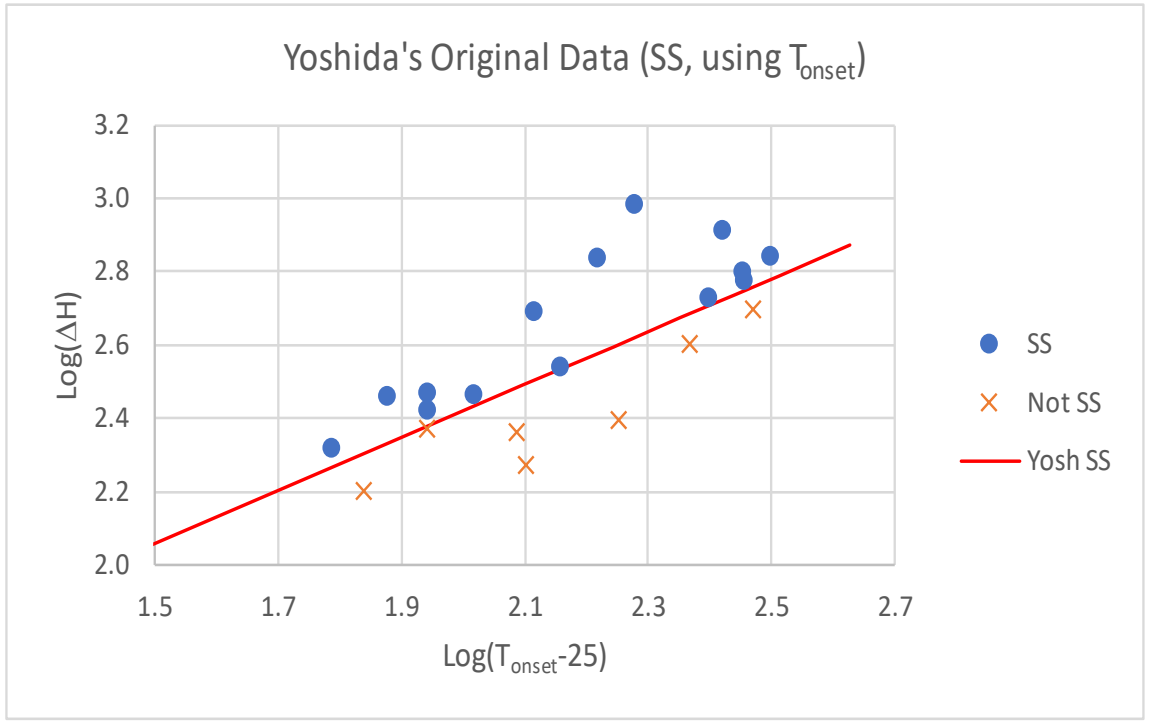
- Journal articles
- Internal databases (Merck, AZ, BI [no structural info])
- Include only chemically-well-defined substances (no mixtures/formulations)
- Attempt to focus on materials of pharmaceutical relevance (intermediates, APIs, reagents) rather than high explosives
- Eliminate materials which have very complex multi-modal DSC behavior
- Not all data is created equal
  - T_{init} vs T_{onset}? Terminology is not always applied consistently
    - For materials where T_{init} cannot be directly estimated from the available data, we developed a correlation to estimate T_{init} given T_{onset}
  - DSC pans: high-pressure sealed bombs?
  - Fallhammer: “positive” vs. “negative” can be subjective and may vary from company to company or user to user
    - Some papers reported only performing a manual “hammer test” with an actual hammer

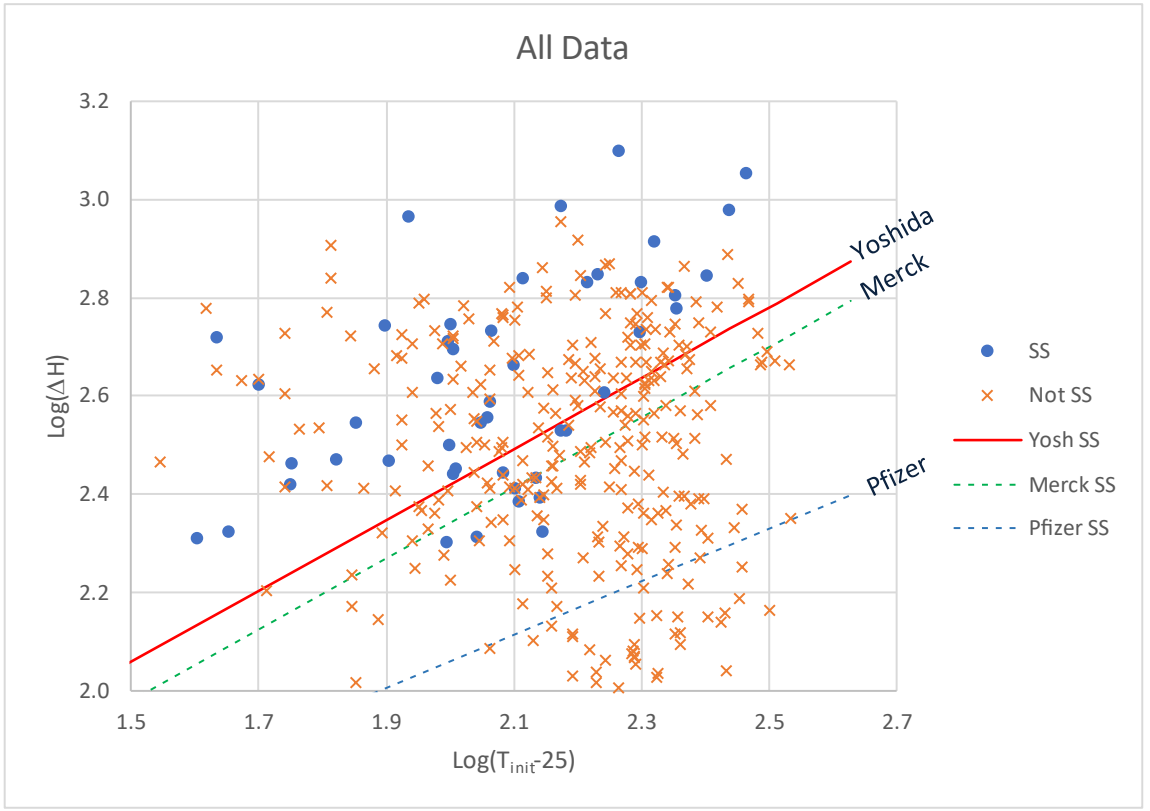
Data Set

- ~380 compounds in the current dataset, ~300 where structural information is available
  - ~70 are shock-sensitive
  - Some classes of compounds of interest include:
    - Yoshida's original ~20 compounds
    - Pharmaceutical intermediates and active pharmaceutical ingredients (APIs)
    - Azodicarboxylates
    - Peptide coupling reagents
    - Diazo compounds
    - Fluoro-azidopyridines
    - Oxidation reagents

Results

Original Yoshida data set vs. our expanded data set





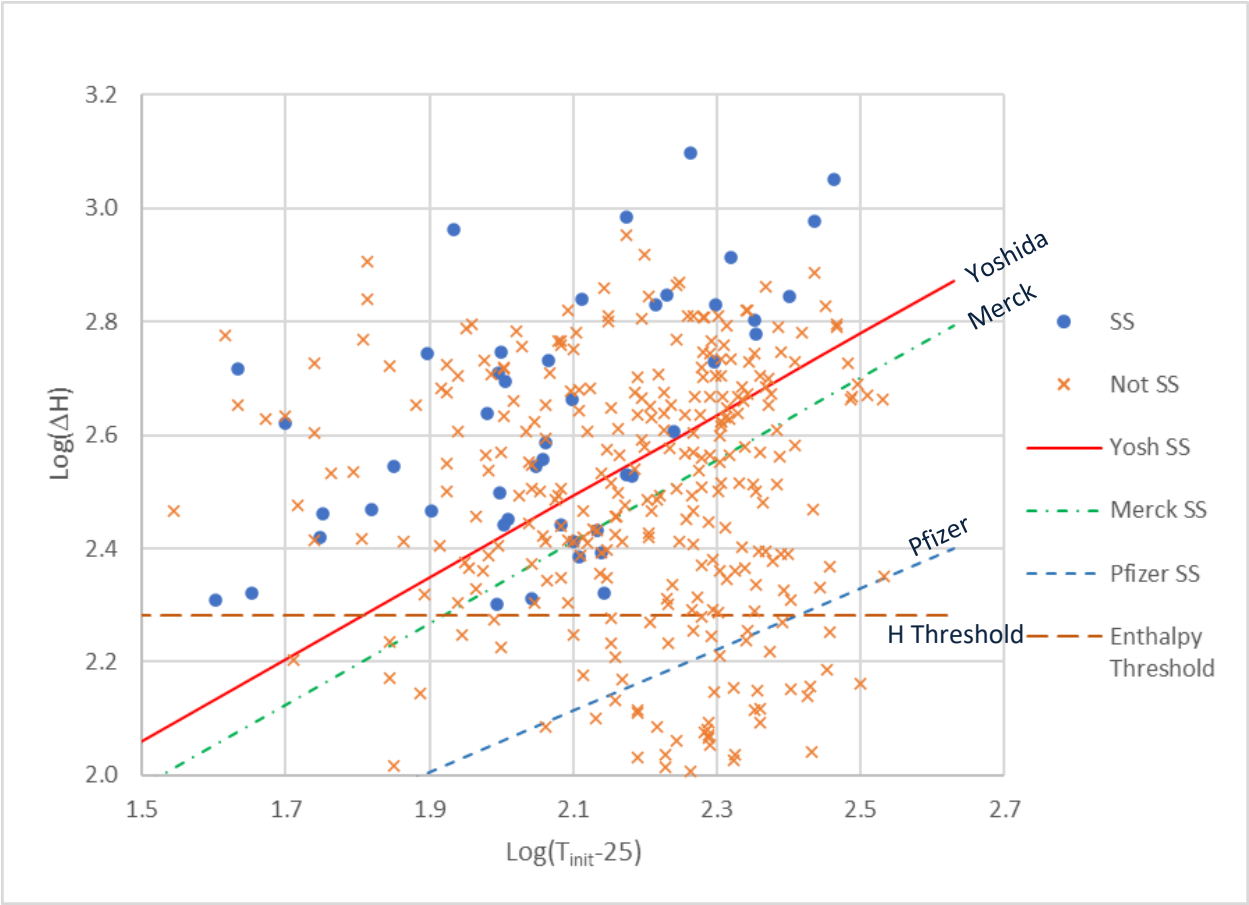
- Some false negatives, many false positives
- Merck and Pfizer modifications reduce false negatives, but greatly increase false positives

- No clear indication that these two variables are suitable predictors!

Results

Accuracy Comparison (full dataset)

	Yoshida	Merck	Pfizer	Enthalpy > 800 J/g
True Pos.	57	62	70	69
True Neg.	179	137	49	67
False Pos.	133	175	263	245
False Neg.	13	8	0	1
Accuracy %	62%	52%	31%	36%



Additional variables analyzed

Besides DSC data, supplement with additional structural information

- MW, molecular formula (# of each type of atom)
  - C, H, N, O, M
- Oxygen Balance and OB100
- # of each HEFG/explosophore
  - Nitro, azo, azide, N-O, N-N, N-S, O-O, C≡C
- Explosophore density (explosophores/MW)

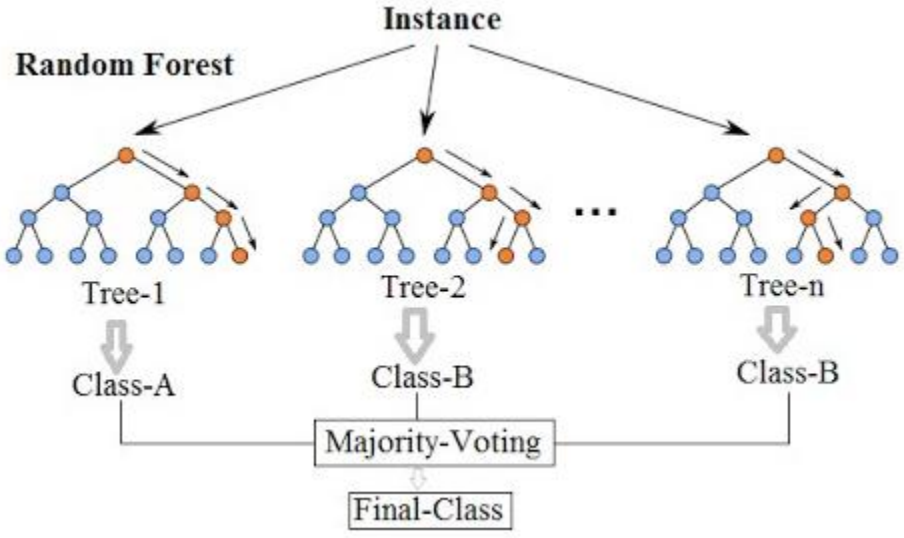
Some of these are inter-correlated, reduce the number of variables by eliminating highly correlated variables

Some of the explosophores show up very infrequently in the dataset

Random Forest approach - methodology

Due to the Dichotomous Nature of the Data (SS or not SS), this is a Classification problem

- Software randomly separates 10% of the datapoints to serve as diagnostic
- Uses the other 90% to generate 1000 different “decision trees” which use the variables to predict whether a given compound is shock-sensitive or not
- Each of the 1000 trees gets a “vote” when making a prediction for an unknown compound
- Final “model” is a “black-box” set of trees, not a simple equation that can be written down

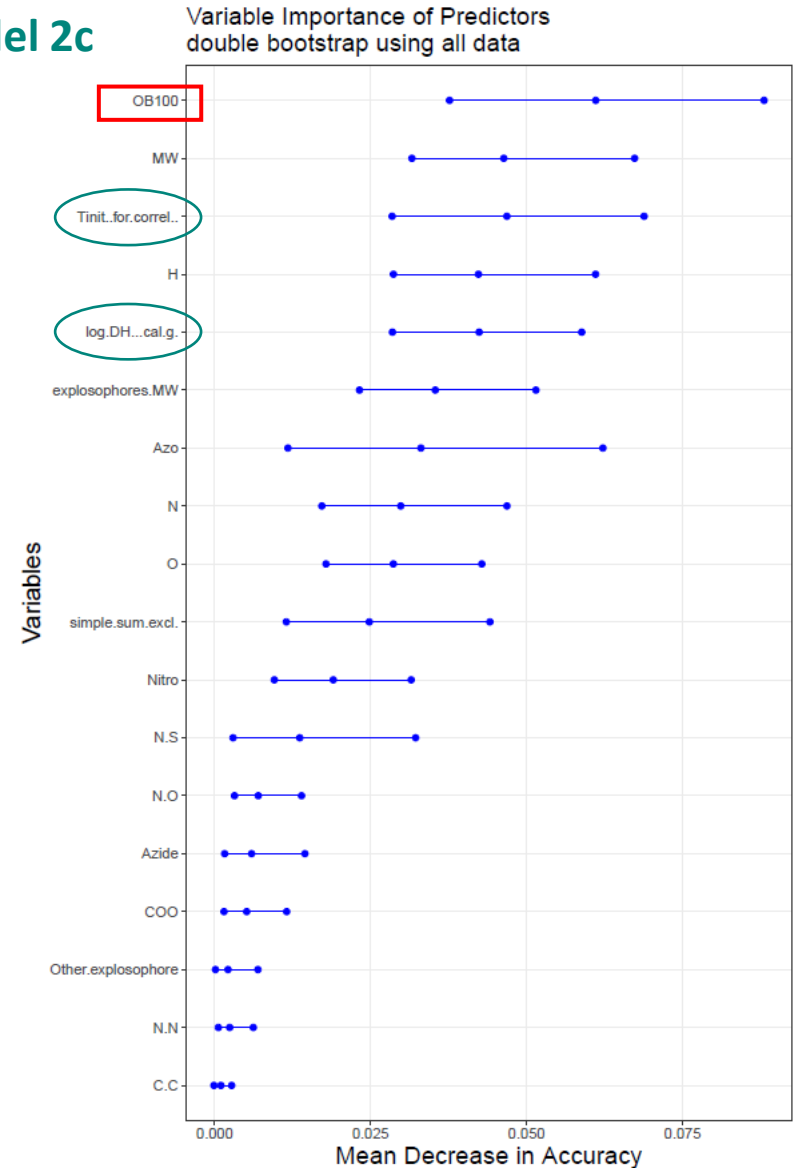


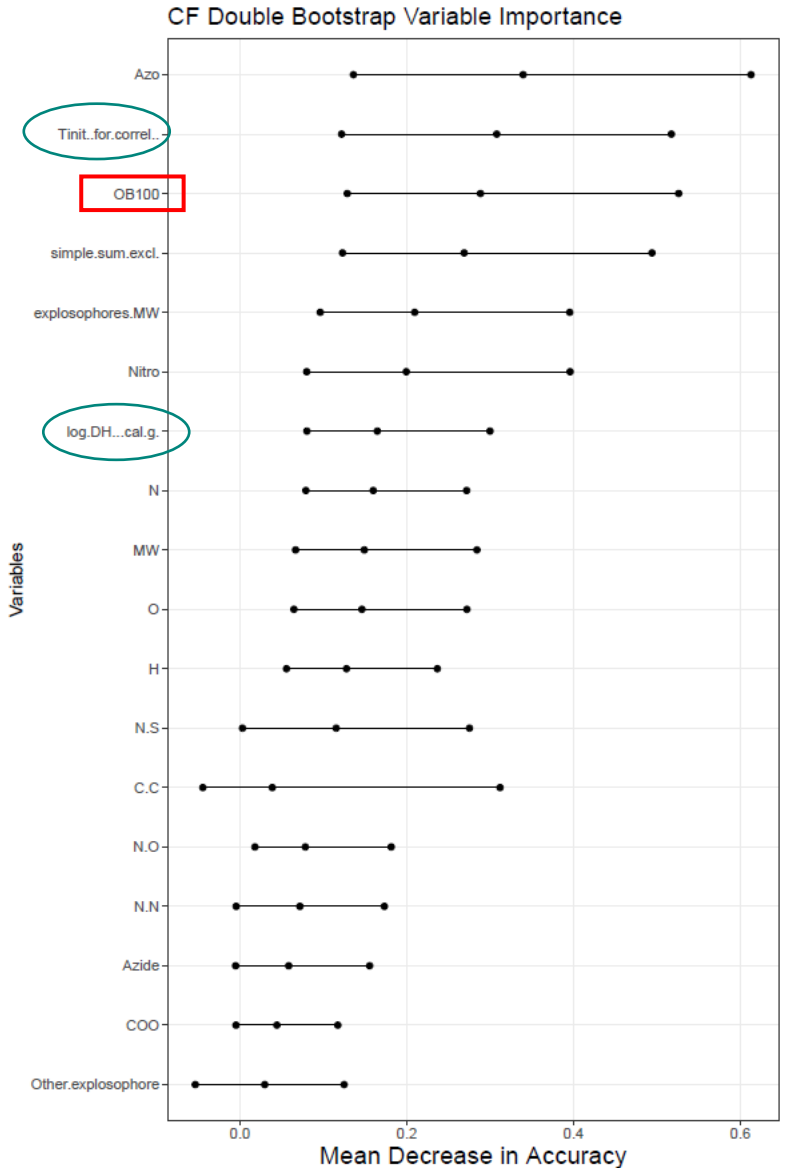
Source:
https://www.nvidia.com/en-us/glossary/random-forest/

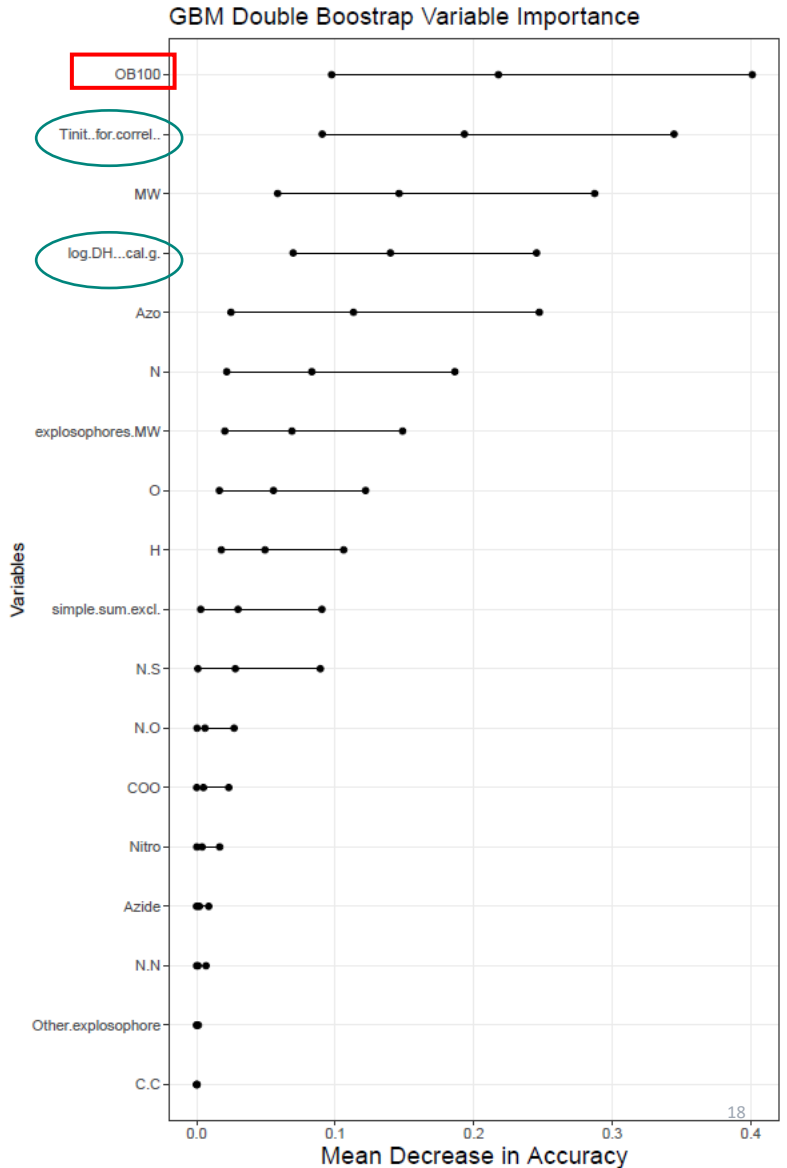
Machine Learning approach – initial results

Yoshida's variables

Model 2c







Logistic Regression Using “Top 5” Variables

Accuracy Comparison (only compounds with structural information)

	Yoshida	Merck	Pfizer	Enthalpy > 800 J/g	Log.Reg 2c (semi-conservative)	Log.Reg 2c (conservative)
True Pos.	57	62	70	69	59	69
True Neg.	97	58	8	20	166	88
False Pos.	127	166	216	204	58	136
False Neg.	13	8	0	1	11	1
Accuracy %	52%	41%	27%	30%	85%	53%

Prob Score = f(OB100, Azo, T_{init}, etc.)

Term	L-R ChiSquare	Prob>Chi Sq
MW	3.33	0.0681
OB100	31.46	<.0001*
Azo	23.60	<.0001*
Tinit	25.94	<.0001*
log(DH)	3.02	0.082
MW*OB100	4.64	0.0312*
MW*Azo	0.67	0.414
MW*Tinit	0.76	0.3841
MW*log(DH)	3.71	0.0541
OB100*Azo	2.61	0.1062
OB100*Tinit	12.84	0.0003*
OB100*log(DH)	5.41	0.0201*
Azo*Tinit	1.56	0.2115
Azo*log(DH)	5.78	0.0162*
Tinit*log(DH)	2.03	0.1544

Conclusions

- Yoshida correlation should be applied very cautiously if at all, as its fundamental underpinnings have not been fully validated
  - The process safety community can do better with modern data analytics
  - The old variables of log(T-25) and log(ΔH) are not by themselves good predictors of SS
- Our work on SS analysis is ongoing (hope to publish later this year)
  - Adding in oxygen balance and a few 2-factor interactions can greatly improve the correlation

Acknowledgements

- **Nelson Afanador** (machine learning and regression analysis)
- **Ben Dobson** (AZ) data compilation
- Andre Janssen (BI)
- Tetsuji Itoh (translation/interpretation of Yoshida's original paper)
- Tao Chen (Fallhammer testing)
- Prof. Jimmie Oxley (URI)
- Athina Kominia (Kodak)
- Shasha Zhang (BMS)