Lignin Carbohydrate Complexes – Learning the Structure-Property Relation

Matthias Stosiek, Joakim Löfgren, Daryna Diment, Mijung Cho, Davide Rigo, Marie Alopaeus, Michael Hummel, Chunlin Xu, Mikhail Balakshin†, Patrick Rinke

FinnCERES Materials Cluster

What is Lignin?

Biomass (birch wood)

Cellulose

Hemicellulose

<u>Lignin:</u>

- Aromatic polymer found in the cell walls of plants.
- Binds cellulose and hemicellulose together in woody plants.
  => high stiffness and resistance to rot

Why is Lignin interesting?

Pulping industry

Paper

Cellulose

Waste Product

<u>Lignin</u>

Abundant underutilized biomaterial

Why is Lignin interesting?

Pulping industry

Paper

Cellulose

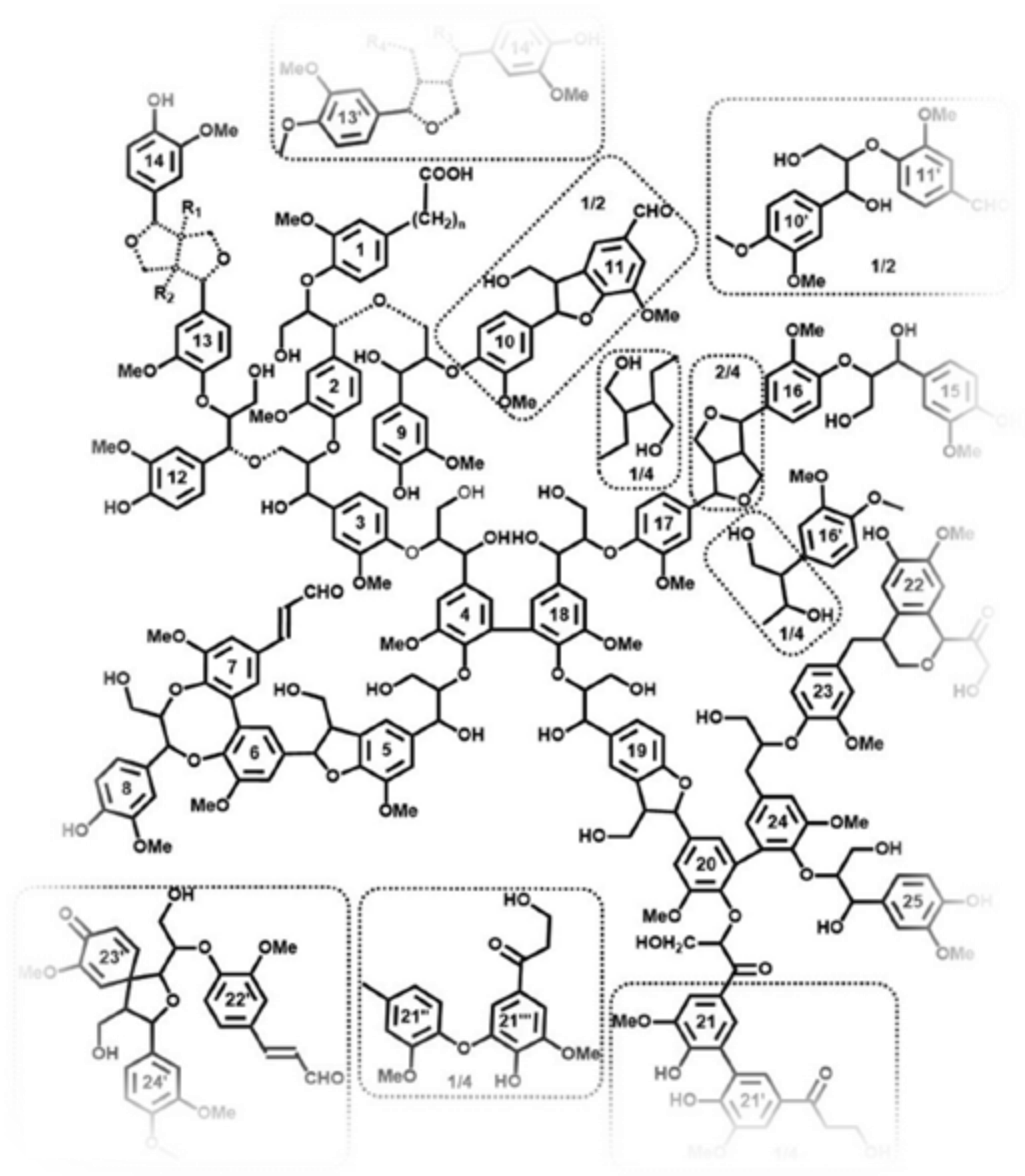
Waste Product

Lignin

Valorization

Lignin Bio-based products

Learning the Structure-Property relationship



- Lignin is chemically very complex. Can form Lignin Carbohydrate Complexes (LCC).
- Lack of chemical knowledge doesn't allow for prediction of LCC properties from structure.
  -> structure-property relationship is essential for lignin tailored for specific applications.

With machine learning, we can deduce the structure-property relationship from experimental data!

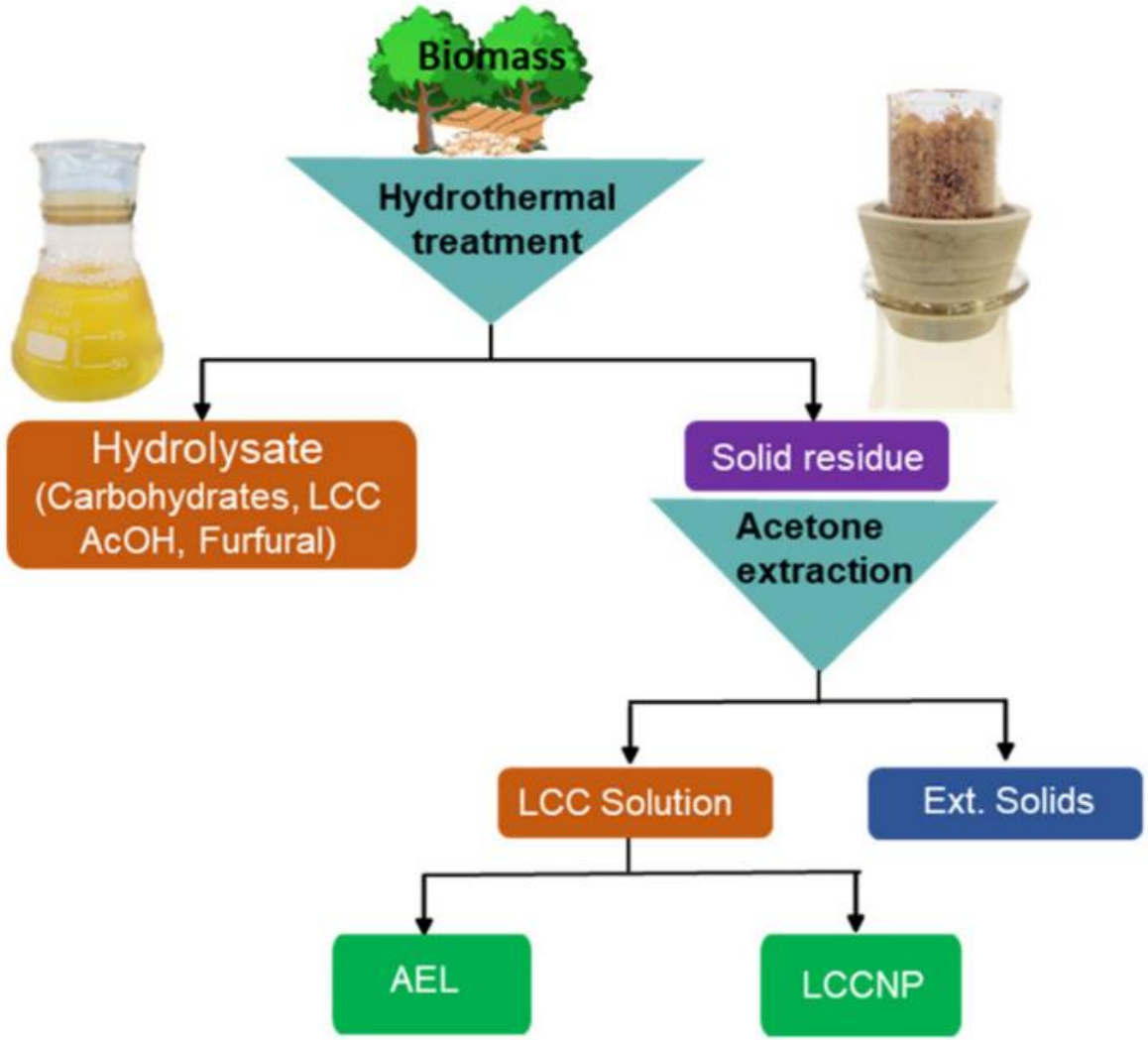
Outline

1. Introduction - Experiments
2. Structure-property relation

Experiments

1. LCC sample preparation: biorefineries
2. LCC structural characterisation: NMR spectroscopy
3. LCC property characterisation: Antioxidant Activity – Radical Scavenging Index

AqSO Lignin biorefinery

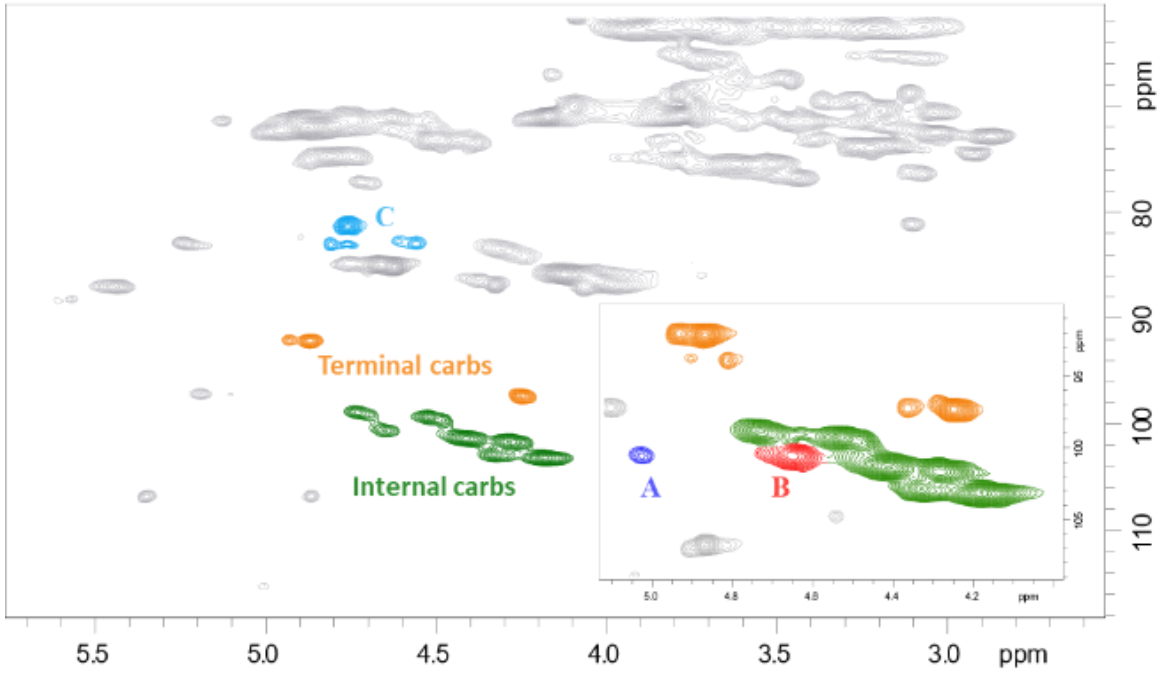


- Highly-tunable, green, industrially scalable process.

Dmitry Tarasov et al., Green Chem., 2022, 24, 6639-6656

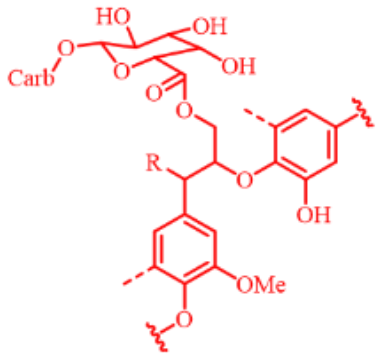
P. Schlee et al., ChemSusChem2023, e202300549

2D Heteronuclear Single Quantum Coherence (HSQC) NMR spectroscopy

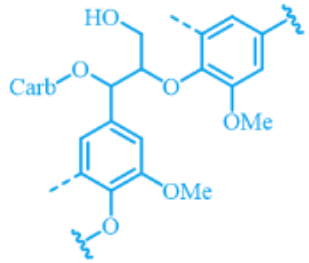


A
PhGly

B
GlcU Ester



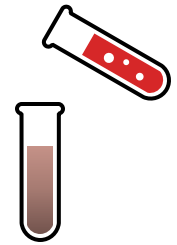
C
Benzyl Ether



Antioxidant Activity – Radical Scavenging Index (RSI)

- Antioxidants are very important as additives in organic substrates such as plastics, fibers, adhesives, oils, …
- Very high antioxidant activity in Lignins has been reported, higher than standard commercial products.

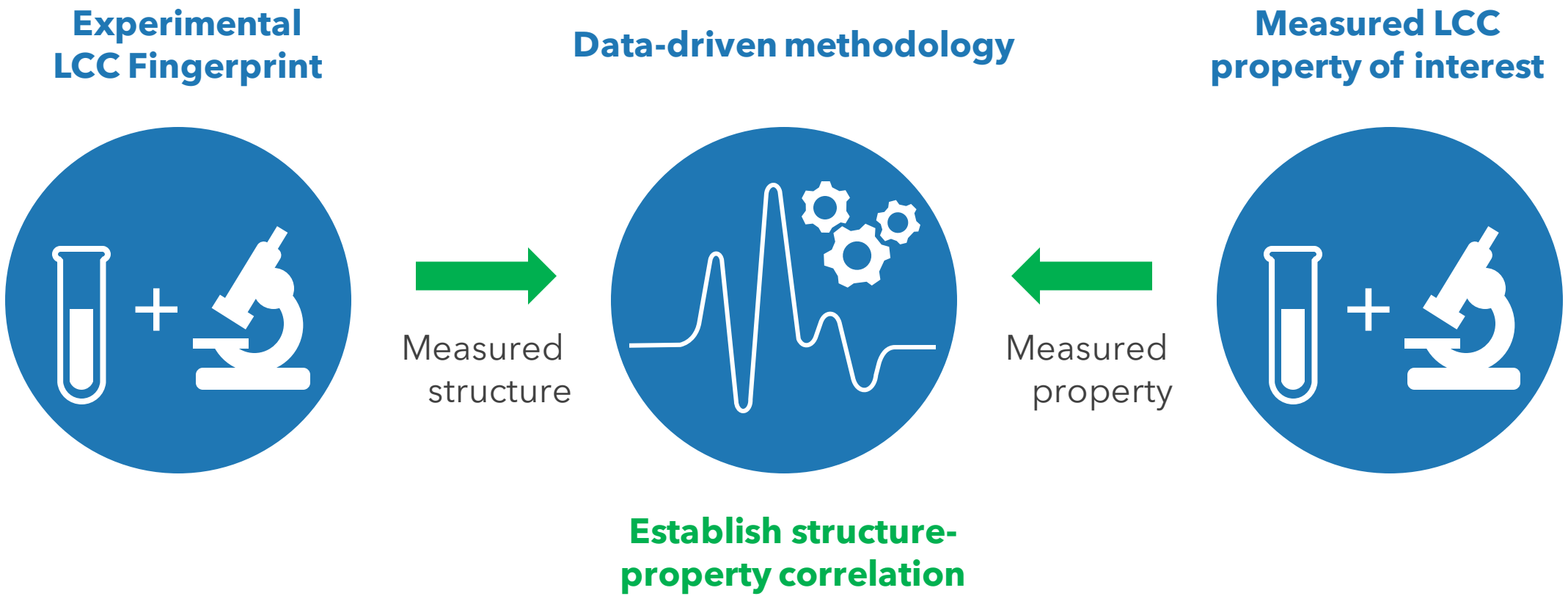
1,1-diphenyl-2-picrylhydrazyl (DPPH)



Lignin

How much DPPH
Is left in steady-state?

Learning the structure-property relationship



Learning the structure-property relationship

Experimental LCC Fingerprint

Data-driven methodology

Measured LCC property of interest

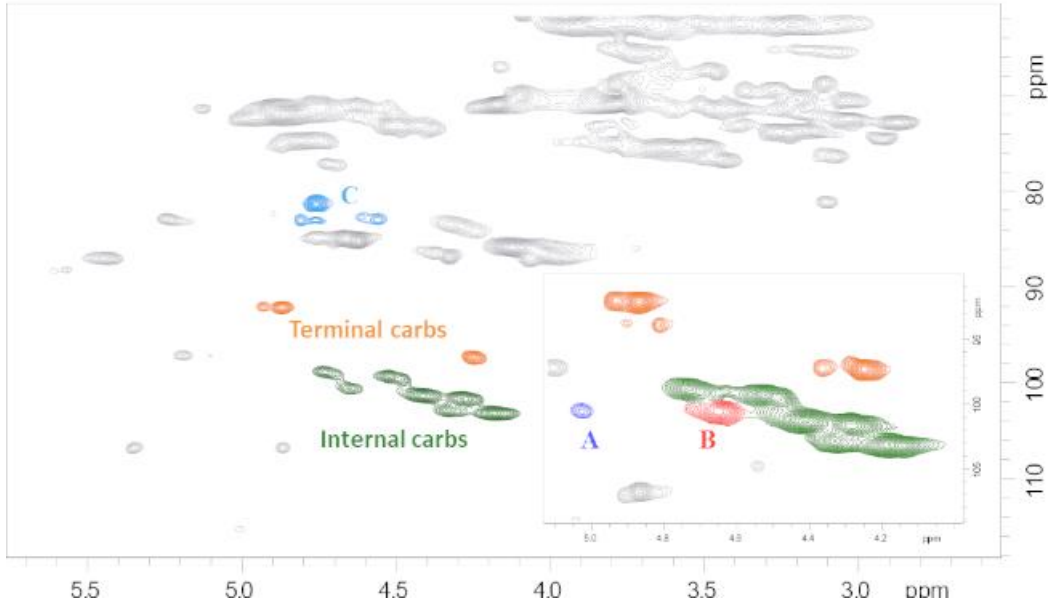
Establish structure-property correlation

Understand which kind of LCC leads to optimal properties

Gain insight on the underlying chemistry

What machine learning model do we use?

We have 62 samples so far, for which both RSI and NMR has been measured.



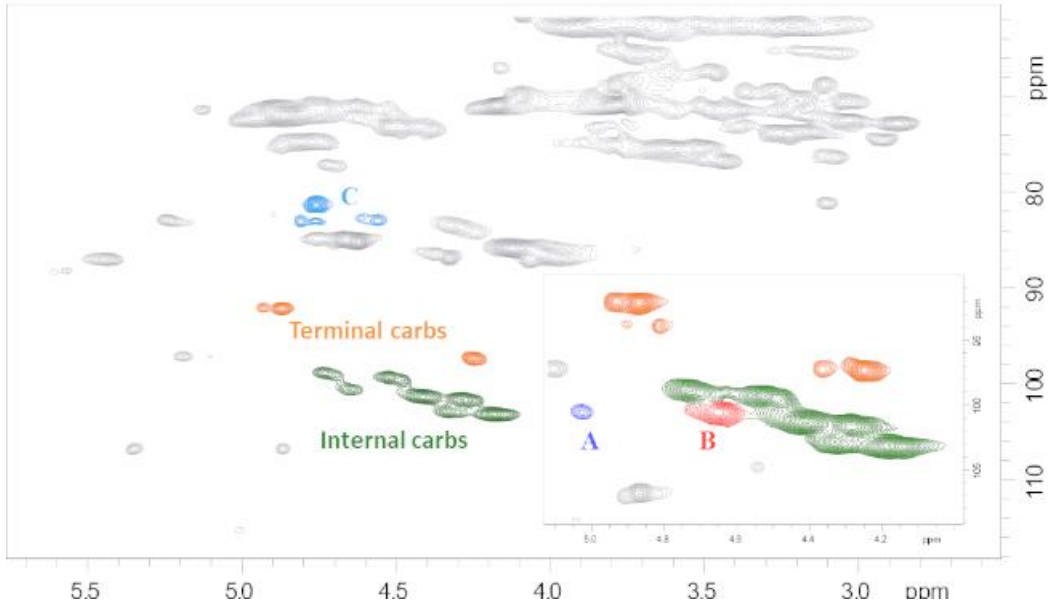
Regression

Measured RSI

- Allows us to predict properties from structure and quantify how important which feature is.

What machine learning model do we use?

We have 62 samples so far, for which both RSI and NMR has been measured.

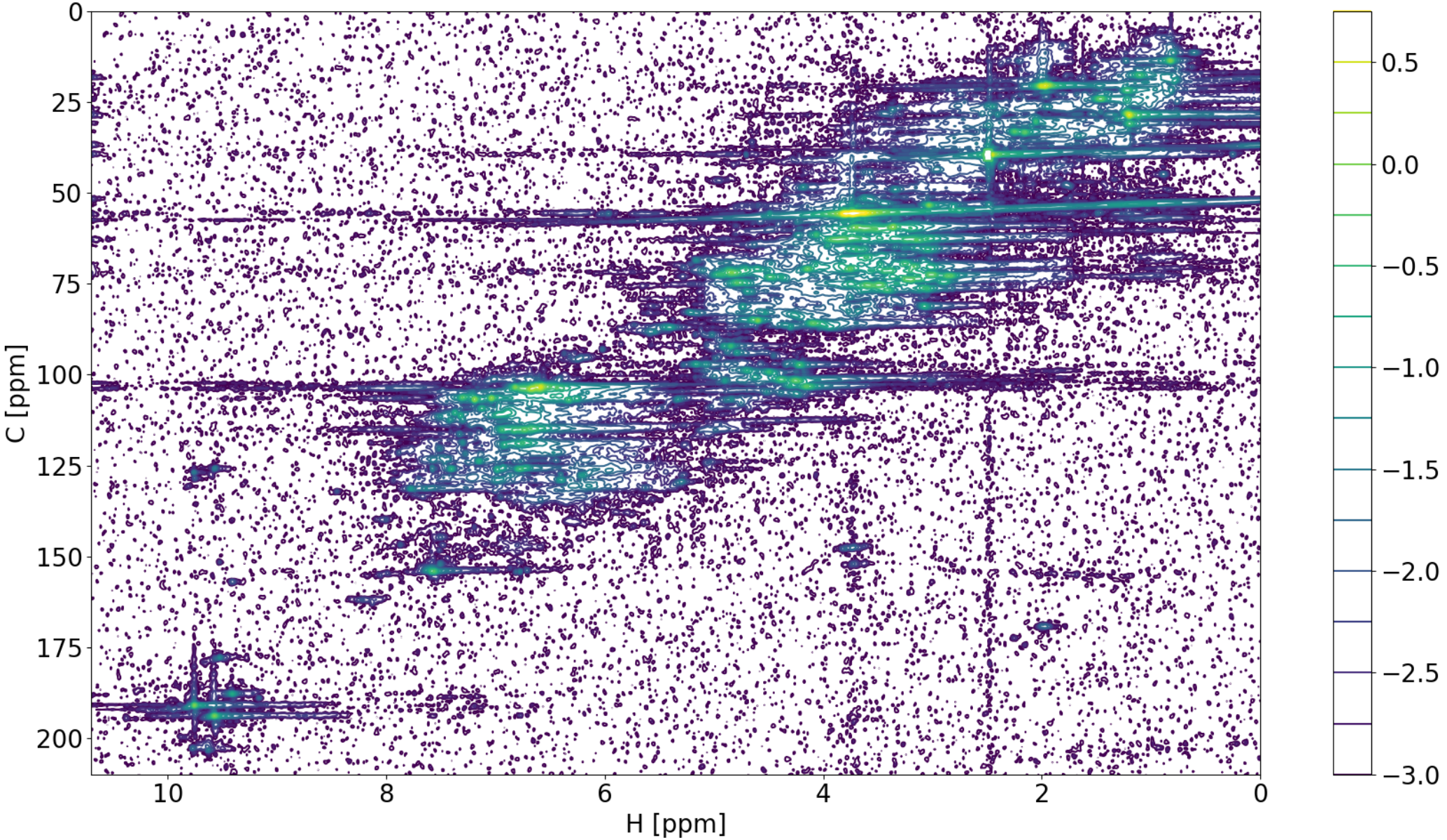


Regression

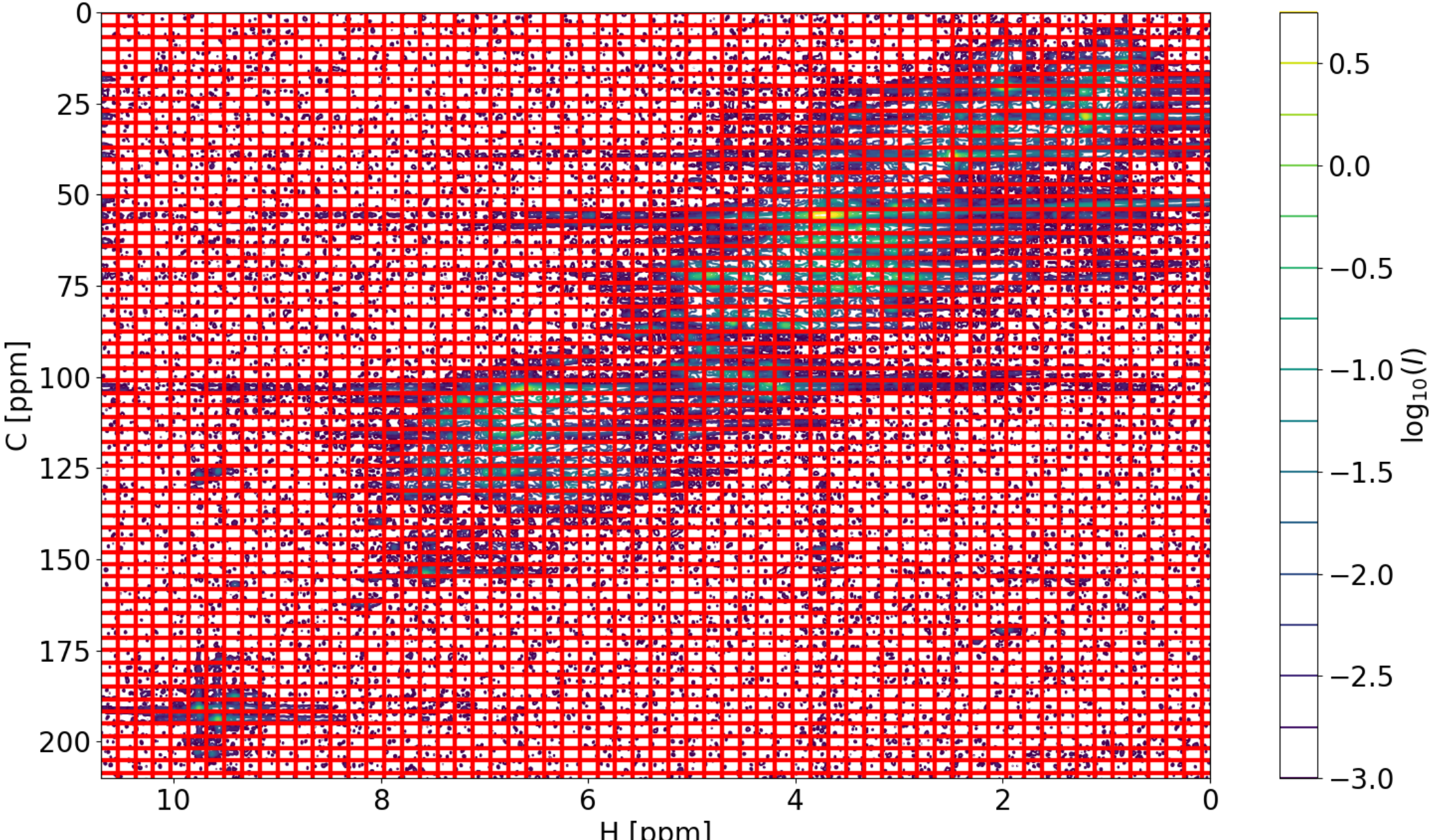
Measured RSI

- Allows us to predict properties from structure and quantify how important which feature is.
- NMR spectrum is too high-dimensional as input!

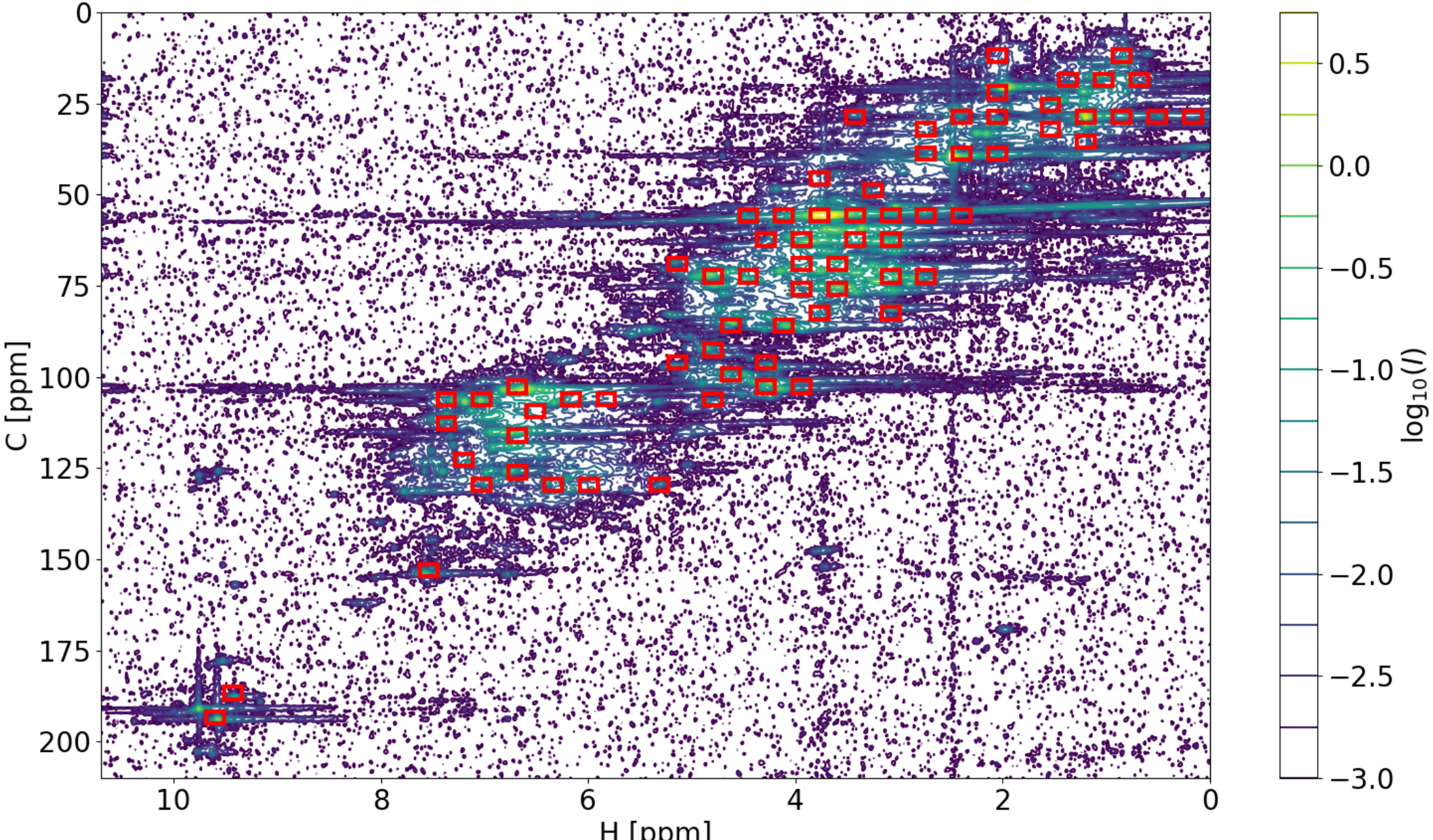
Example Spectrum



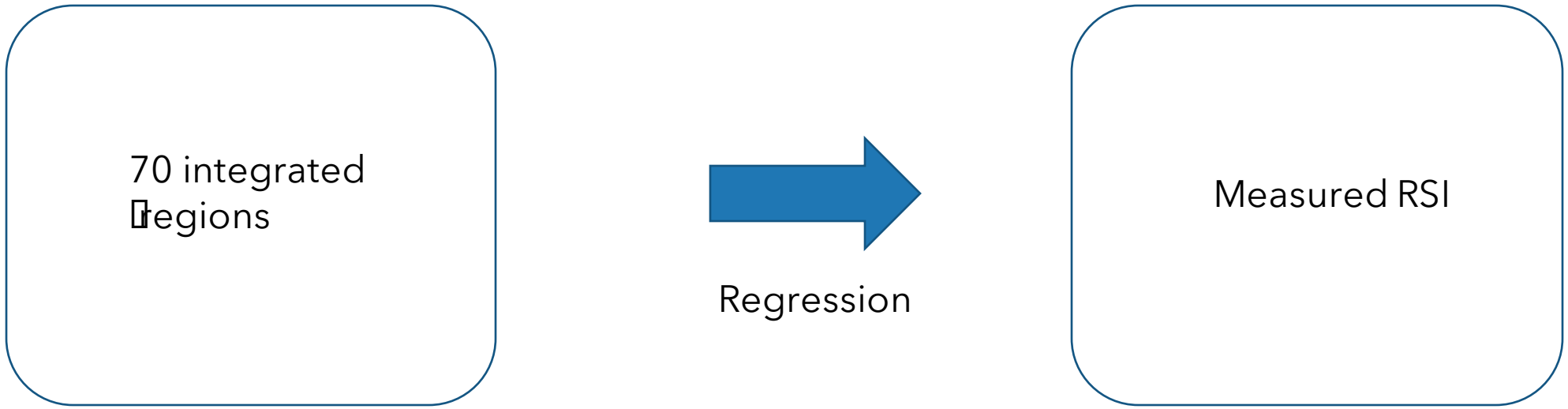
Integrated Regions



Integrated Regions

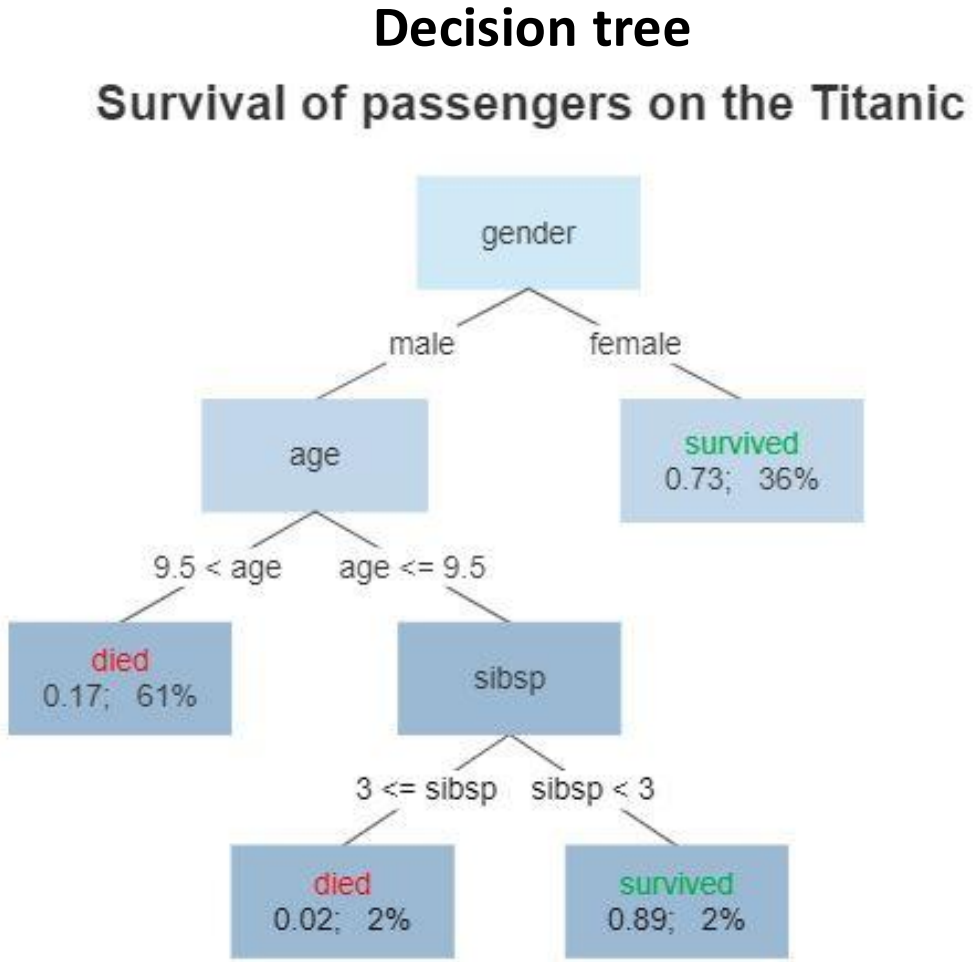


Property Prediction



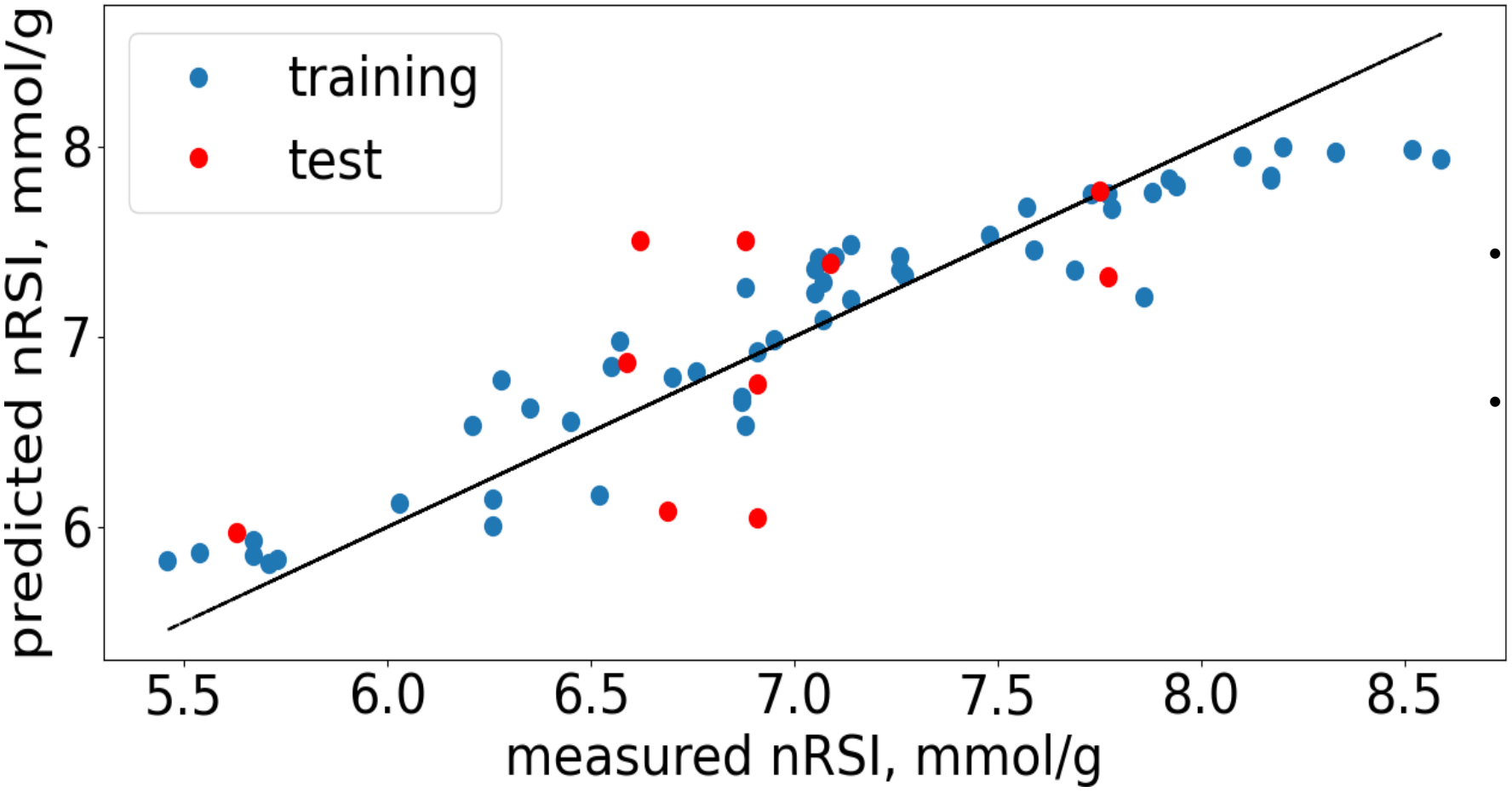
Random Forest Regression

Decision tree



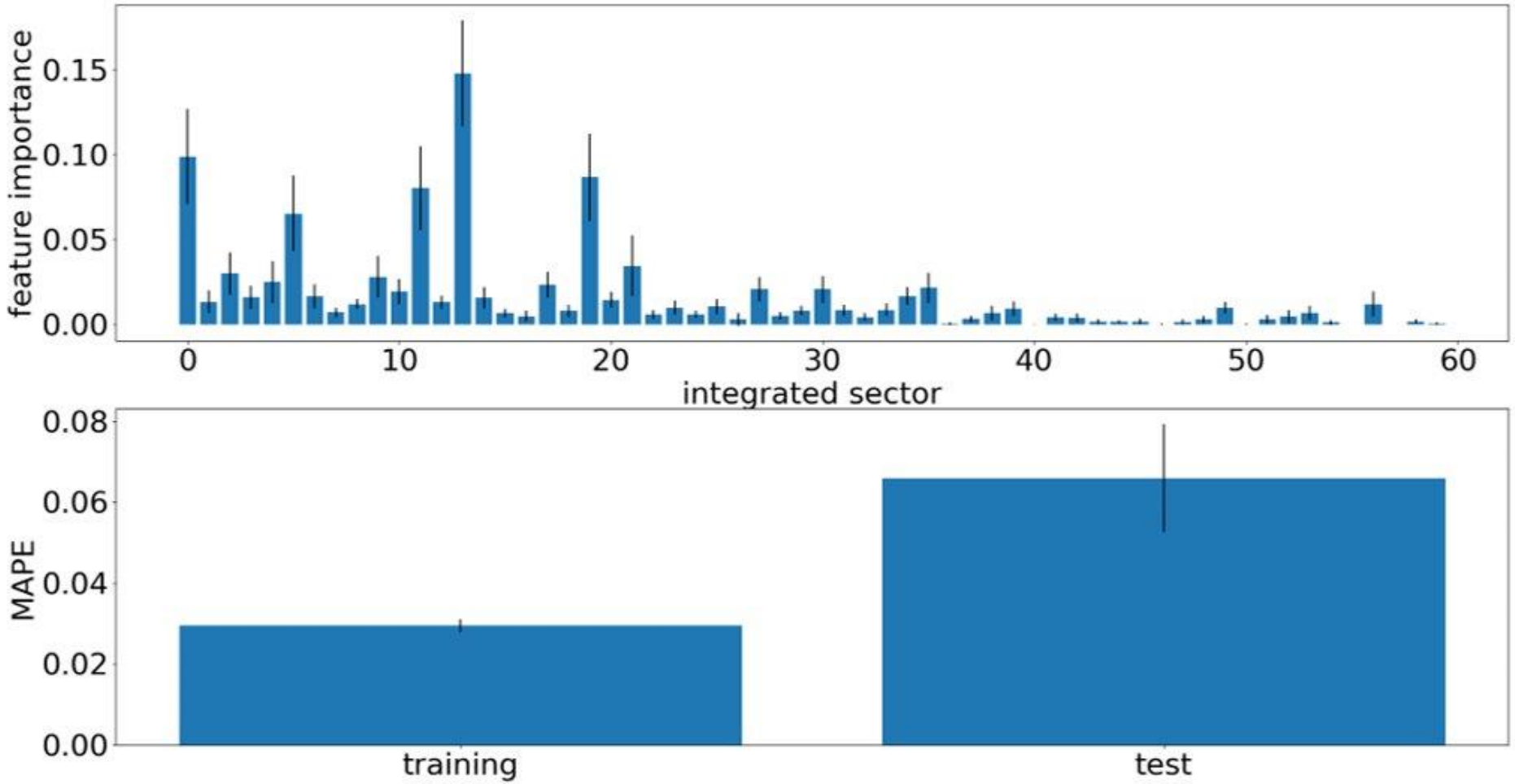
- In random forest regression, we build many decision trees and the model prediction is the average prediction of the all trees.
- Each decision tree is created with a subset of the data using only a subset of the features for each split.
- Random forests work quite well, when data is scarce.
- Feature importance analysis is possible.

RSI regression

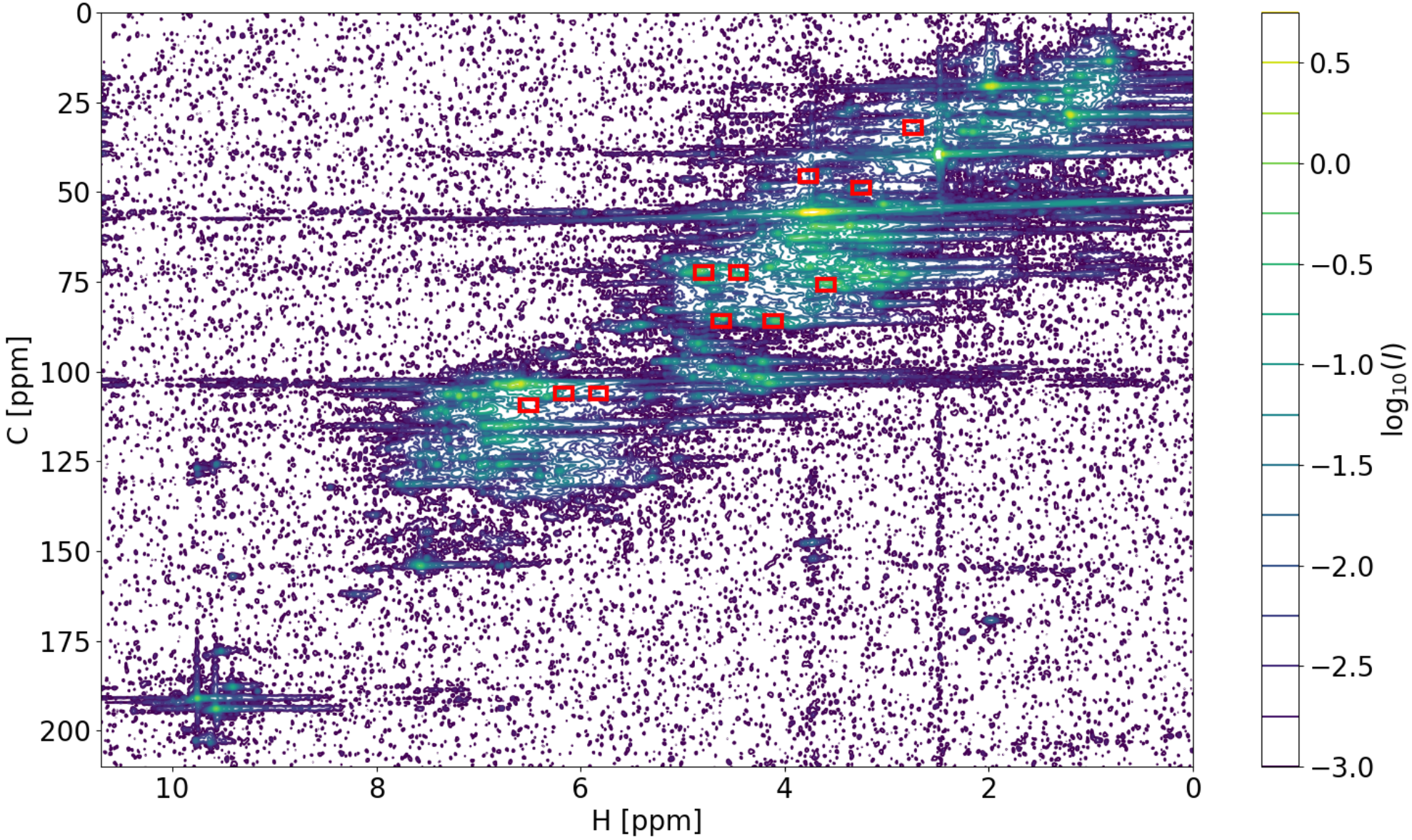


- MAPE of training data is 3.3% and of test data is 6.7%.
- We get a meaningful prediction from RF model!

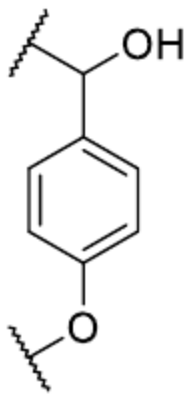
Feature Importance Analysis



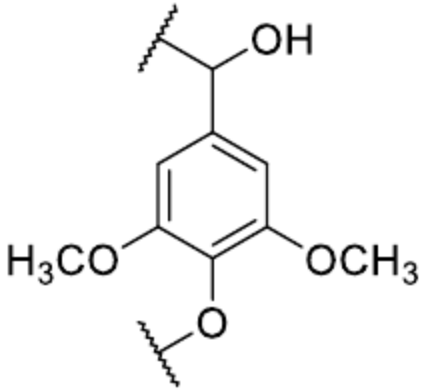
Integrated Regions – Feature Importance Filtered



Chemical Analysis

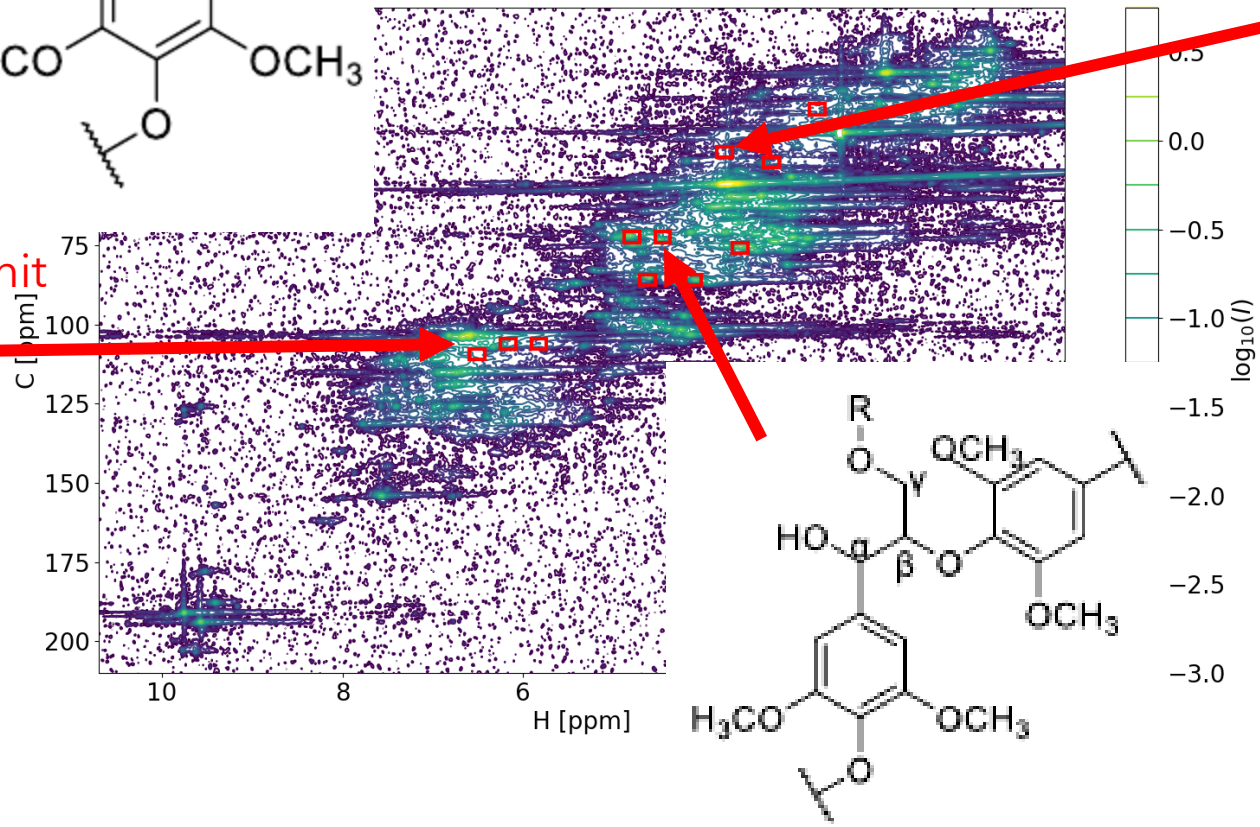


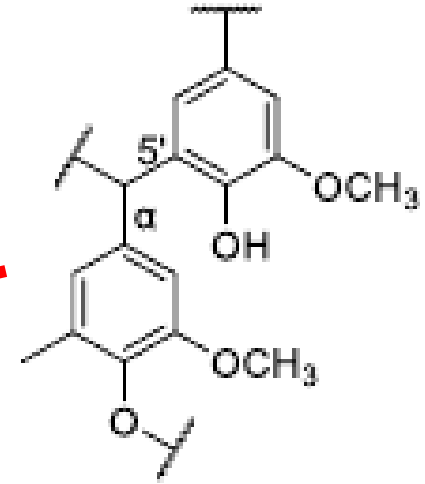
G-Unit



S-Unit

Aromatics





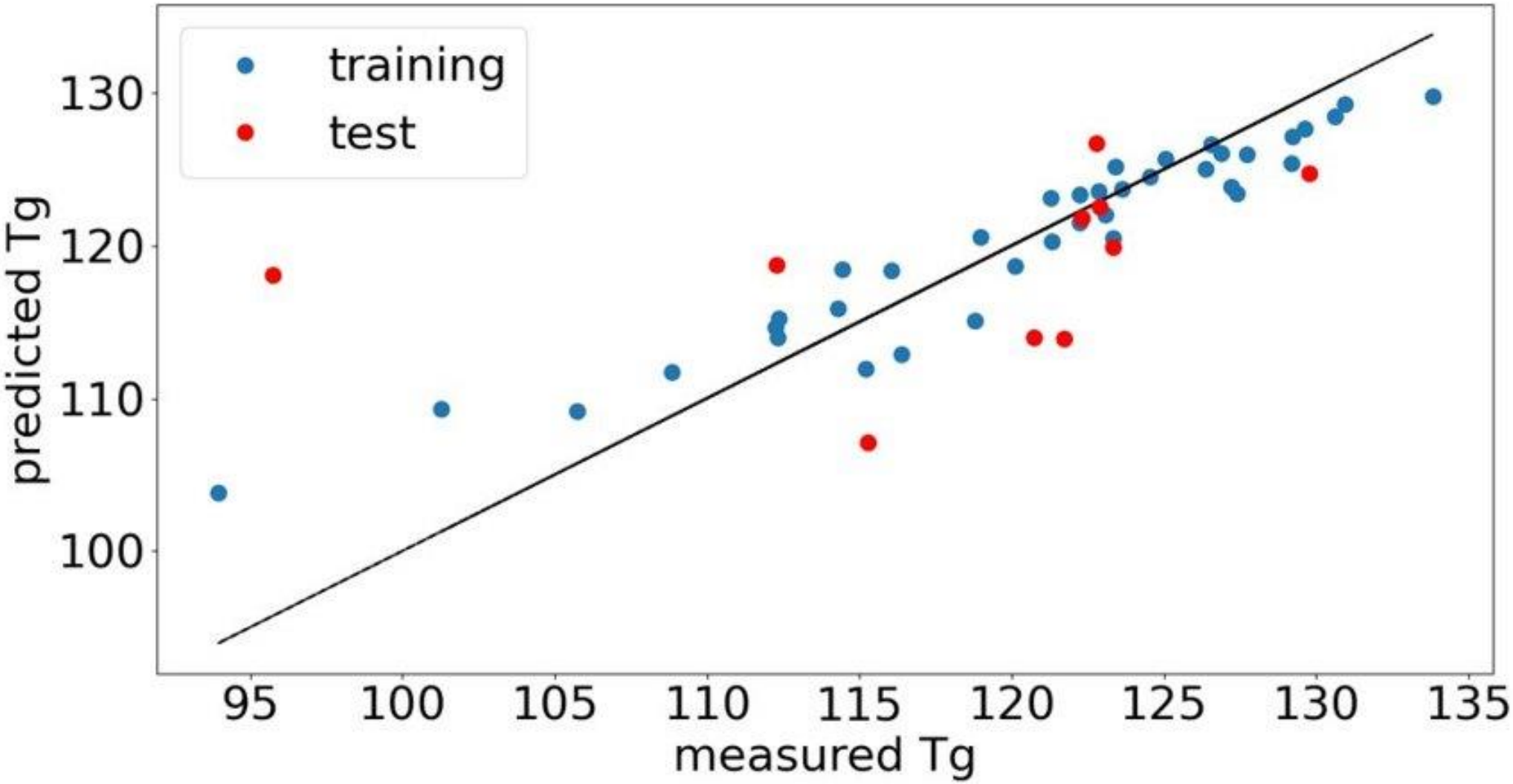
α-5'

β-O-4

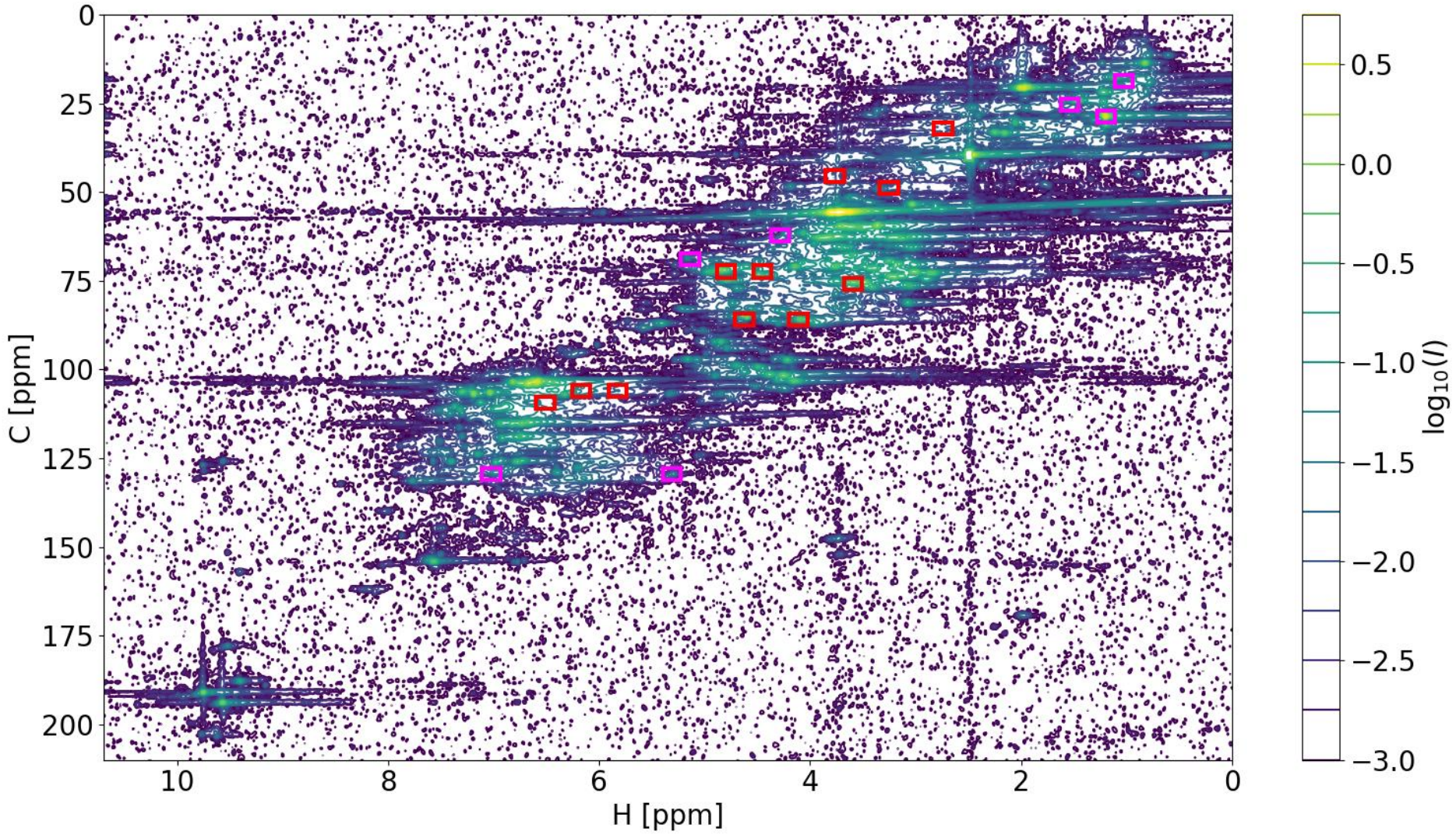
Other properties

1. Glass transition temperature Tg
2. Molecular weight Mw
3. Thermal degradation metrics
4. Surface Tension

Model Performance Tg



Important Sectors Tg



Conclusion

- Prediction of RSI (and other properties) from NMR spectra is possible.
- With a feature importance analysis, we can determine how much each structural group determines each property.
- In discussion with our experimental collaborators we are working on the chemical interpretation.

Thank you for your attention!