

Σύγκριση δεδομένων HPLC και COD

Βιδιάνος Γιαννίτσης

December 7, 2023

Contents

1	Dependencies	2
2	Data Reading and Processing	2
3	Plotting	3
4	Αποτελέσματα για το πείραμα 10/11	4
5	Αποτελέσματα για το πείραμα 28/11	4
6	Πείραμα 23/10	5
6.1	HPLC Data	5
6.2	COD Data	6
6.3	Plotting	6
6.4	Τελικά code blocks	6

Abstract

Η ανάλυση COD είναι μία ανάλυση που μετράει την συνολική οργανική ύλη μίας ένωσης με βάση το πόσο οξυγόνο θέλει για να οξειδωθεί πλήρως. Η ανάλυση της HPLC μετράει την συγκέντρωση για συγκεκριμένα χημικά είδη. Συγκρίνοντας τα δεδομένα των δύο αναλύσεων μπορούμε να δούμε πόση από την οργανική ύλη που μπορεί να υπάρχει μετράει όντως η HPLC και πόση είναι άλλες, άγνωστες ενώσεις. Βέβαια, λόγω των μεγάλων αραιώσεων που απαιτούνται για να μετρηθεί το COD, πολλές φορές δεν έχει ικανοποιητική ακρίβεια, το οποίο μπορεί να δημιουργήσει προβλήματα στην σύγκριση αυτή. Το αρχείο αυτό αποτελεί ένα literate document που επεξηγεί την διαδικασία αυτή αναλυτικά και δείχνει τα αποτελέσματα της, ενώ μπορεί να κάνει tangle τον κώδικα που αναφέρεται στα κατάλληλα scripts ή να τον κάνει weave σε ένα pdf. Αξίζει να αναφερθεί πως γίνεται έντονη named code blocks τα οποία βοηθάνε στο να τρέξουμε πολλές φορές το ίδιο code block με βάση του noweb syntax που προσφέρει το org-mode. Καθώς το όνομα του code block δεν γίνεται weaved, θα υπάρχει με bold το όνομα πριν ακριβώς από κάθε code block, ώστε να μπορεί να κατανοηθεί και τότε κάποιο γίνεται inserted σε άλλα.

1 Dependencies

Καθώς το project αυτό είναι δομημένο με το πακέτο DrWatson της Julia για να κάνει facilitate reproducibility, πρέπει σε όλα τα αρχεία να υπάρχουν τα lines που ενεργοποιούν το DrWatson και πηγαίνουν στο κατάλληλο project. Έπειτα, είναι επίσης απαραίτητο να κάνουμε include κάποια functions για την ανάλυση των δεδομένων το οποίο υπάρχει στο src directory και είναι amply documented εκεί.

dependencies

```
using DrWatson
@quickactivate "Masters_Thesis"
include(sourcedir("cod_balance.jl"))
include(sourcedir("filenames.jl"))
```

2 Data Reading and Processing

Αρχικά πρέπει να διαβάσουμε τα δεδομένα των δύο αναλύσεων από τα CSV τους. Θα υποθέσουμε πως τα δεδομένα της HPLC είναι ήδη επεξεργασμένα και απλώς θα διαβάσουμε τα csv με τις συγκεντρώσεις. Για να παραχθούν αυτά, μπορούμε να κάνουμε refer στο αντίστοιχο notebook, στο οποίο εξηγείται και περισσότερο τι πειράματα έχουν γίνει και πως τα διαχειριζόμαστε. Το COD είναι έτοιμο στην processed μορφή του για τα πειράματα στις 10 και 28 Νοεμβρίου ενώ για το αρχικό πείραμα της κινητικής, υπάρχει μόνο ένας πίνακας με διάφορες πληροφορίες συμπεριλαμβανομένου της απορρόφησης του COD, οπότε απαιτείται ένα παραπάνω processing step εκεί.

Αρχικά θα γραφεί το code block που κάνει το generic processing των δεδομένων και βασίζεται στο variable date και μετά θα γραφούν τα blocks που το τρέχουν για τις διαφορετικές ημερομηνίες. Εκτός από να διαβάσουμε τα δεδομένα της HPLC, πρέπει να κάνουμε extract μόνο την πρώτη και την τελευταία μέρα επειδή αυτές υπάρχουν στο COD. Έπειτα, για να έχουμε συγκρίσιμα μεγέθη, υπολογίζουμε το COD-eq κάθε ένωσης και κάνουμε τις μετατροπές. Αυτό φαίνεται παρακάτω.

cod_{dataextraction}

```
mix_amount = ["0", "1", "2", "4", "8"]
df0_conc = CSV.read(get_conc_csv(date, mix_amount[1]), DataFrame)
df1_conc = CSV.read(get_conc_csv(date, mix_amount[2]), DataFrame)
df2_conc = CSV.read(get_conc_csv(date, mix_amount[3]), DataFrame)
df4_conc = CSV.read(get_conc_csv(date, mix_amount[4]), DataFrame)
df8_conc = CSV.read(get_conc_csv(date, mix_amount[5]), DataFrame)

cod_meas = CSV.read(get_cod_csv(date), DataFrame)

df0_conc0 = Vector{Float64}(df0_conc[1, 2:8])
df1_conc0 = Vector{Float64}(df1_conc[1, 2:8])
df2_conc0 = Vector{Float64}(df2_conc[1, 2:8])
df4_conc0 = Vector{Float64}(df4_conc[1, 2:8])
df8_conc0 = Vector{Float64}(df8_conc[1, 2:8])

df0_conc72 = Vector{Float64}(df0_conc[4, 2:8])
df1_conc72 = Vector{Float64}(df1_conc[4, 2:8])
```

```

df2_conc72 = Vector(df2_conc[4, 2:8])
df4_conc72 = Vector(df4_conc[4, 2:8])
df8_conc72 = Vector(df8_conc[4, 2:8])

cod_yields = [cod_sucrose(), cod_glucose(), cod_fructose(), cod_acetate(),
               cod_propionate(), cod_lactate(), cod_ethanol()]

df0_cod0 = sum(df0_conc0.*cod_yields)
df1_cod0 = sum(df1_conc0.*cod_yields)
df2_cod0 = sum(df2_conc0.*cod_yields)
df4_cod0 = sum(df4_conc0.*cod_yields)
df8_cod0 = sum(df8_conc0.*cod_yields)
cod0_theor = [df0_cod0, df1_cod0, df2_cod0, df4_cod0, df8_cod0]
cod0_meas = cod_meas.COD_0./1000
cod0_error = cod0_meas - cod0_theor

df0_cod72 = sum(df0_conc72.*cod_yields)
df1_cod72 = sum(df1_conc72.*cod_yields)
df2_cod72 = sum(df2_conc72.*cod_yields)
df4_cod72 = sum(df4_conc72.*cod_yields)
df8_cod72 = sum(df8_conc72.*cod_yields)
cod72_theor = [df0_cod72, df1_cod72, df2_cod72, df4_cod72, df8_cod72]
cod72_meas = cod_meas.COD_72./1000
cod72_error = cod72_meas - cod72_theor

```

3 Plotting

Με αυτά τα δεδομένα, μπορούμε να κάνουμε τα συγκριτικά plots μεταξύ των δύο.

```

codcompplots

label = ["COD Measurement" "HPLC Measurement"]
plot_type = "comparison"

cod0_plot = groupedbar([0, 1, 2, 3, 4], [cod0_meas cod0_theor],
                        label = label, xlabel = "Amount of mix (ml)",
                        ylabel = "COD (g/l)", legend =:bottom,
                        title = "COD comparison t=0",
                        xticks = (0:4, mix_amount))
savefig(cod0_plot, get_plot_name("cod_0", date, plot_type))

cod72_plot = groupedbar([0, 1, 2, 3, 4], [cod72_meas cod72_theor],
                        label = label, xlabel = "Amount of mix (ml)",
                        ylabel = "COD (g/l)", legend =:bottom,
                        title = "COD comparison t=72 h",
                        xticks = (0:4, mix_amount))
savefig(cod72_plot, get_plot_name("cod_72", date, plot_type))

cod_plot = plot(cod0_plot, cod72_plot, layout = (2,1), size = (900, 600))

```

```
savefig(cod_plot, get_plot_name("cod", date, plot_type))
```

4 Αποτελέσματα για το πείραμα 10/11

Έχοντας τα παραπάνω code blocks, μπορούμε ορίζοντας το date και τρέχοντας τα να παράγουμε τα απαιτούμενα plots. Το συγκεντρωτικό plot φαίνεται παρακάτω. Επίσης, αυτό ακριβώς είναι και αυτό που θέλουμε να γίνει tangled, μόλις φυσικά κάνουμε tangle τα dependencies.

```
<<dependencies>>
```

```
date = "10_11"  
<<cod_data_extraction>>  
<<cod_comp_plots>>
```

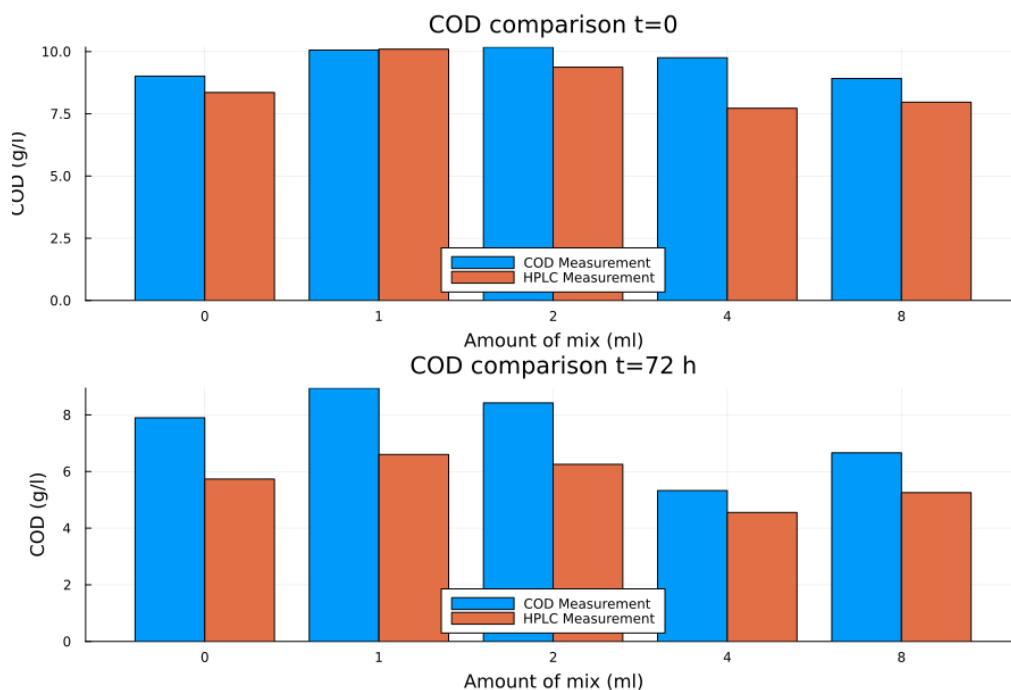


Figure 1: Σύγκριση της μέτρησης COD άμεσα ή μέσω της HPLC - Πείραμα 10/11

5 Αποτελέσματα για το πείραμα 28/11

Ομοίως, μπορούμε αλλάζοντας το date να τρέξουμε και το πείραμα της 28/11 χωρίς καμία πρακτική αλλαγή στον κώδικα.

```
<<dependencies>>
```

```
date = "28_11"
<<cod_data_extraction>>
<<cod_comp_plots>>
```

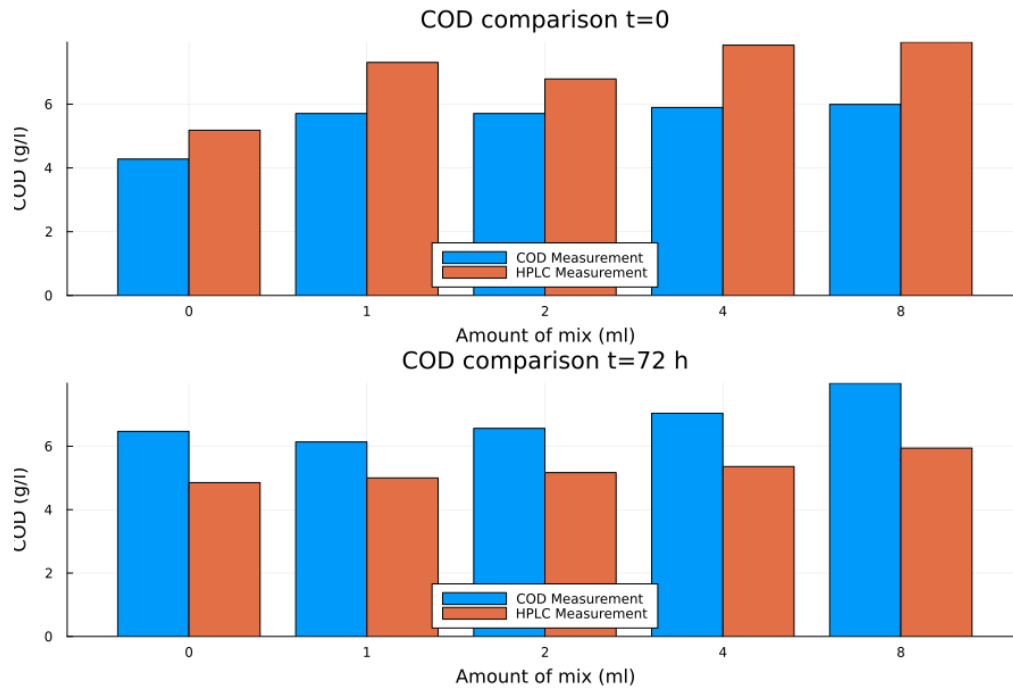


Figure 2: Σύγκριση της μέτρησης COD άμεσα ή μέσω της HPLC - Πείραμα 28/11

6 Πείραμα 23/10

Στο πείραμα αυτό υπήρχε καθημερινή μέτρηση του COD, εκτός από την τελευταία μέρα που ήταν μετά το σαββατοκύριακο οπότε έχει νόημα να γίνει σύγκριση σε όλα τα διαθέσιμα σημεία.

6.1 HPLC Data

Αρχικά διαβάζουμε τα δεδομένα της HPLC και κάνουμε τις μετατροπές σε COD-eq.

```
hplcdataprocessing2310
```

```
date = "23_10"
df1_conc = CSV.read(get_conc_csv(date, "1"), DataFrame)
t1 = df1_conc[1:19, 1]
df1_conc = df1_conc[1:19, 2:8]
df2_conc = CSV.read(get_conc_csv(date, "2"), DataFrame)
t2 = df2_conc[1:15, 1]
df2_conc = df2_conc[1:15, 2:8]

cod_yields = [cod_sucrose(), cod_glucose(), cod_fructose(), cod_acetate(),
               cod_propionate(), cod_lactate(), cod_ethanol()]
```

```
cod_1_theor = [sum(Vector(df1_conc[i, :]).*cod_yields) for i in 1:19]
cod_2_theor = [sum(Vector(df2_conc[i, :]).*cod_yields) for i in 1:15]
```

6.2 COD Data

Έπειτα, κάνουμε την επεξεργασία του COD data, το οποίο για αυτό το πείραμα δεν υπήρχε διαθέσιμο σε συγκέντρωση αλλά μόνο σε απορρόφηση. Όμως ο κώδικας είναι πολύ compact λόγω του helper function `process_cod_data` που κάνει την μετατροπή μόνο του.

```
coddataprocessing2310
```

```
cod_1_meas = process_cod_data("1", date; dilution = 50)
cod_1_error = cod_1_meas - cod_1_theor
cod_2_meas = process_cod_data("2", date; dilution = 50)
cod_2_error = cod_2_meas - cod_2_theor
```

6.3 Plotting

Τέλος, μπορούν να γίνουν τα plots για αυτό το πείραμα που λόγω του πλήθους των στοιχείων έχει νόημα μόνο να είναι σε scatter plot.

```
codbalanceplots2310
```

```
label = ["COD Measurement" "HPLC Measurement"]
plot_type = "comparison"

cod_1_plot = scatter(t1, [cod_1_meas cod_1_theor], label=label,
                    xlabel = "Time (h)", ylabel = "COD (g/l)",
                    title = "Sample 1", markersize = 6)
savefig(cod_1_plot, get_plot_name("cod_1", date, plot_type))

cod_2_plot = scatter(t2, [cod_2_meas cod_2_theor], label=label,
                    xlabel = "Time (h)", ylabel = "COD (g/l)",
                    title = "Sample 2", markersize = 6)
savefig(cod_2_plot, get_plot_name("cod_2", date, plot_type))

cod_plot = plot(cod_1_plot, cod_2_plot, layout = (2,1), size = (900, 600))
savefig(cod_plot, get_plot_name("cod", date, plot_type))
```

6.4 Τελικά code blocks

Έπειτα, βάζουμε το τελικό code block που με χρήση του noweb syntax θα κάνει tangle στα script files.

```
<<dependencies>>
```

```
<<hplc_data_processing_23_10>>
```

<<cod_data_processing_23_10>>
<<cod_balance_plots_23_10>>

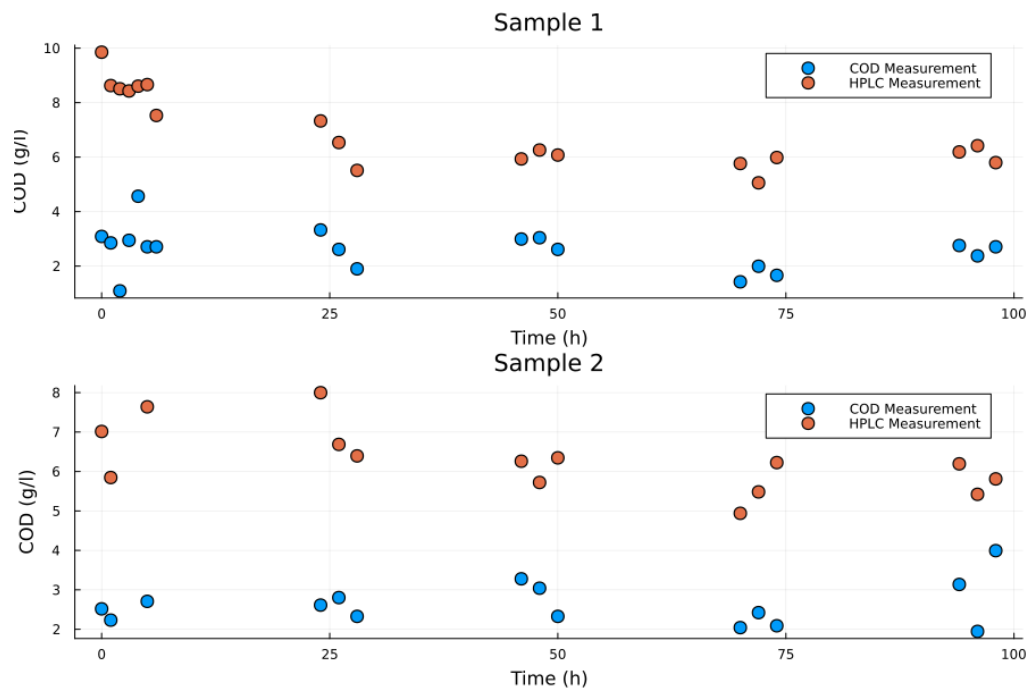


Figure 3: Σύγκριση της μέτρησης COD άμεσα ή μέσω της HPLC - Πείραμα 23/10