



Chapter 4

4. Results and Discussion

This chapter is a detailed discussion of the observation and output obtained from following the methodologies mentioned in Chapter 3. The chapter is divided into 4 subheadings of the 4 objectives. The first findings include the changes in soaking, steaming, and drying processes occurring at different conditions and optimal parboiling parameters for best rehydration property. The second objective's findings include the developed ML models for process analysis of soaking, steaming, and drying using spectral data. The third objective gives us inferences on how storage and ageing affect its cooking quality and physicochemical changes like compositional changes, pasting behavior, textural parameters, molecular bond change, and surface morphology of aged and unaged rice. The temperature-based study also gave us an idea about how temperature affects the physicochemical characteristics of *Komal Chaul*. The fourth objective gives us an NIR-based tool for predicting the degree of ageing using ML models.

4.1 Results and discussion of the kinetic study of the parboiling process and spectral calibration

The first findings include the changes in soaking, steaming, and drying processes occurring at different conditions and the diffusion of rehydration kinetics of produced *Komal Chaul*, in addition to spectral calibration for moisture content during soaking, degree of gelatinization during steaming, and moisture content during drying.

4.1.1 Soaking Kinetics

Water diffusion kinetics of brown rice obtained from dehusking of *Chokuwa* paddy were studied at three temperatures 40, 50, and 60 °C. The moisture content against each time was measured using the gravimetric oven method. The moisture content of rice at several hydration temperatures is shown in Fig. 4.1.1 as a function of soaking time. The plots gave an idea about the dependency of time on the water absorption during soaking at different temperatures. The 3 temperature plots were observed to follow the same pattern; initially, there was a rapid increase in moisture content, later followed by a slow rise in moisture content. Soaking takes place because of the penetration of water molecules into the micellar structure of rice. The rapid uptake of water initially at the first-time interval occurs due to surface tension and due to the porous structure of rice grains. Earlier reports on soaking of brown rice had presented

arguments on the increase of moisture content with an increase in temperature, citing that at higher temperatures partial gelatinization of the rice granules takes place and the rice granules start to swell which results in the high uptake of water at higher temperatures. Increased water uptake with increasing soaking temperature was associated with changes in grain diffusion resistance [1].

4.1.1.1 Peleg equation

The hydration behavior (Fig. 4.1.1) was studied using Peleg's equation mentioned in Eq. 3.4 in Chapter 3. Peleg's equation (Eq. 3.4) represented well the water absorption behavior of brown rice during the soaking process at temperatures 40, 50, and 60°C. The constants k_1 and k_2 were calculated using Eq. 3.4. The slope of $t/M_t - M_0$ against t gave us k_2 and the intercept gave us the value of k_1 (Fig. 4.1.2). The R^2 values greater than 0.92 (Table 4.1.1) confirms the adequacy of the equation for describing the water absorption kinetics of soaking. The Peleg's constants k_1 and k_2 were a function of temperature and decreased with an increase in soaking temperature. k_1 values were inversely related to temperature which indicates the increasing of water absorption rate at higher temperatures. As a result, k_1 and temperature were represented using the Arrhenius equation (Eq. 3.5), which thus solves the estimation of activation energy. The activation energy of the soaking process was calculated to be 42.28 kJ/mol, which doesn't contradict earlier findings of activation energy for soaking brown rice [2].

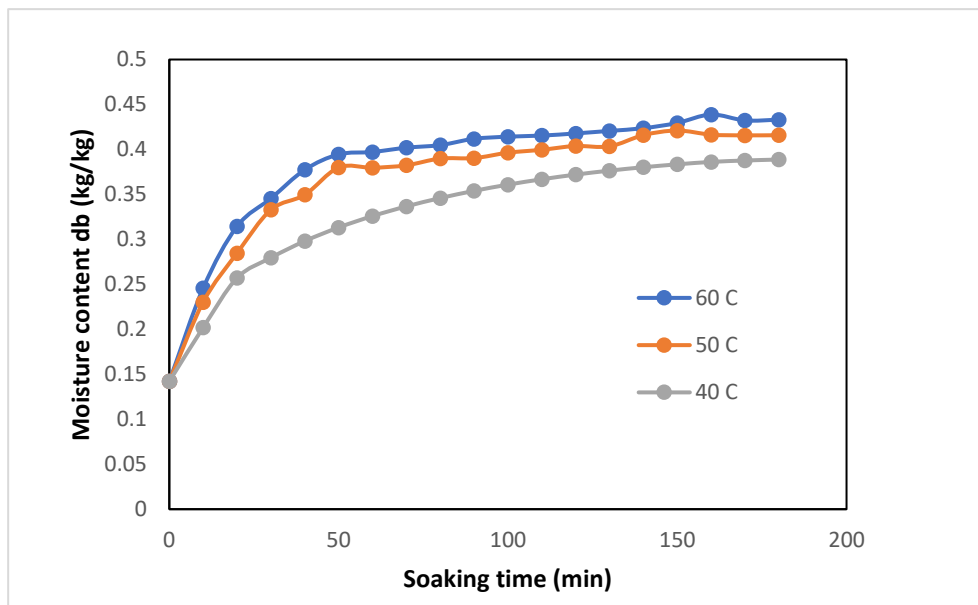


Fig. 4.1.1: Hydration behavior of de-husked *Chokuwa* rice at various temperature

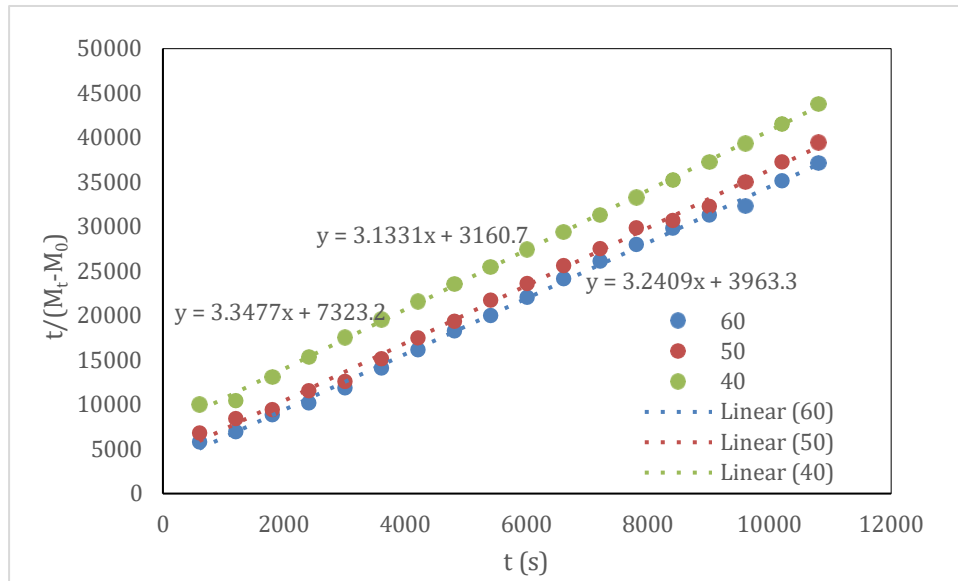


Fig. 4.1.2: Peleg constants estimation plots

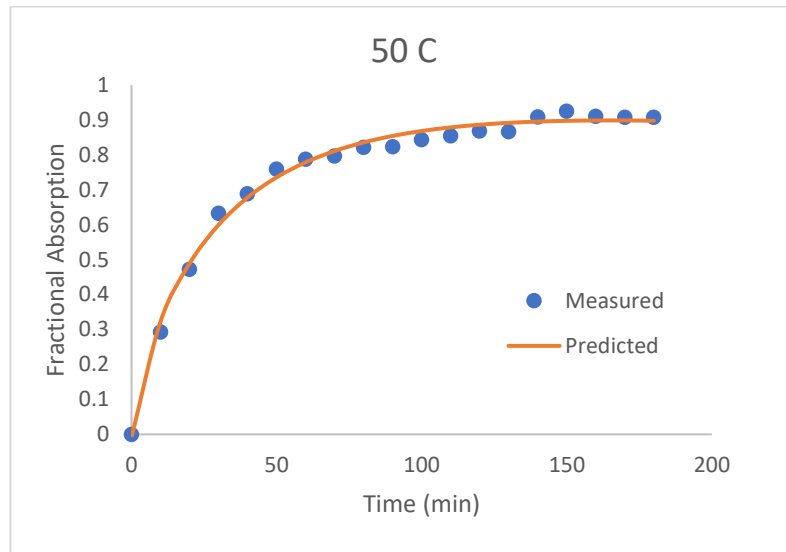
Table 4.1.1: Peleg's constant for the soaking kinetics data at 40, 50, and 60 °C

T°C	$k_1(1/s\%db)$	$k_2(1/\%db)$	R^2	RMSE (fraction)
40	0.043	0.033	0.922	0.120
50	0.041	0.026	0.932	0.112
60	0.039	0.021	0.951	0.038

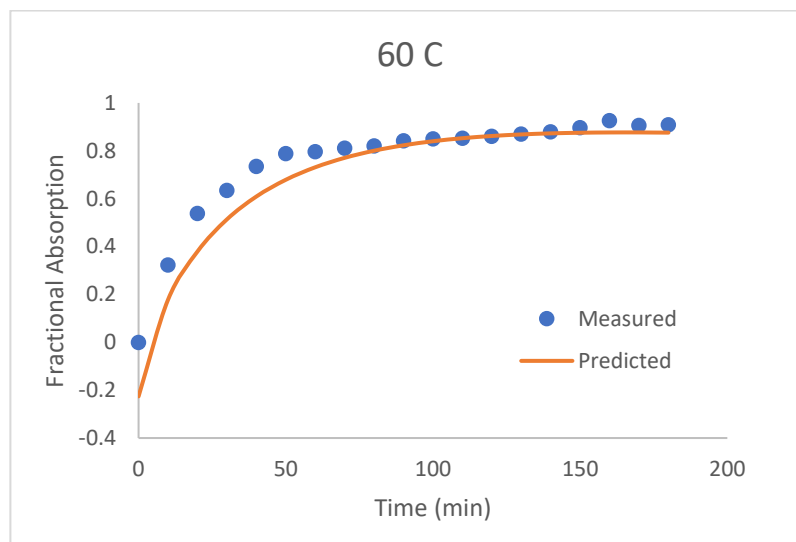
The estimated soaking time from the predicted model for attaining 30% moisture content (wb) at 40, 50, and 60 °C were 185, 135, and 100 min respectively. Wahengbam et al. [3] studied the hydration behavior during brown rice soaking for the *Komal Chaul*, and described the process as temperature dependent, where higher saturation of water is obtained at higher soaking temperature. Previous experiments on the hydration behavior of dehusked *Chokuwa* rice have found similar hydration behavior [3].

4.1.1.2 Generalized Midilli Kucuk equation

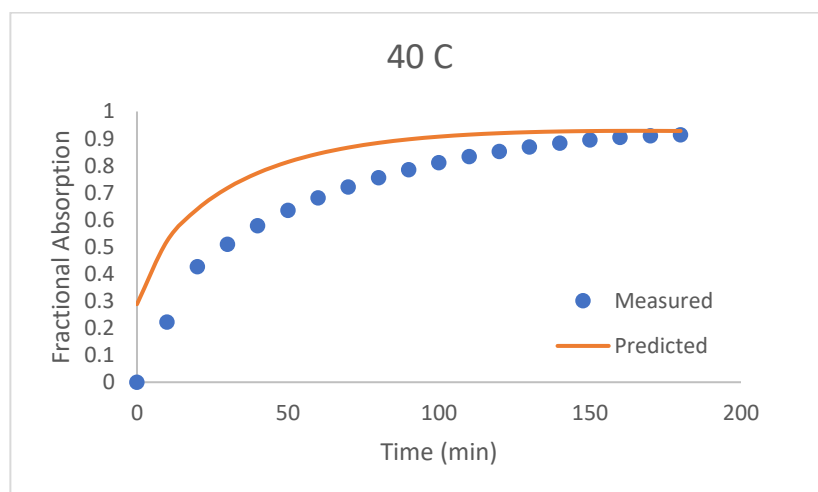
The Midilli-Kucuk model (Eq. 3.8) was developed into a generalized soaking model by multiplication of the time shift factor. Coefficients for the equation were determined by fitting curves to temperature data (40, 50, and 60 °C). The fit for the soaking moisture ratio at all temperatures was highly accurate, with R^2 values exceeding 0.99. The predicted model's root mean square error (RMSE) values indicated improved accuracy (Table 4.1.2) presents the values for the exponential coefficient (a), rate constant (k), power factor (n), and linear constant.



(a) 50°C



(b) 60°C



(c) 40°C

Fig. 4.1.3: Predicted values of moisture using developed Midili Kucuk equation for soaking

To establish a reference equation, as recommended by Bezbaruah and Hazarika [4], the developed model was adjusted accordingly. The rate constant (k) increased with rising temperature, while the power factor of time (n) showed a higher value than the rest of the two temperatures. This can be attributed to the exponential relationship between the soaking rate constant and temperature [4]. By applying a time shift calculation (using Eq. 3.9), the master curve, represented by the equation for the central condition (50 °C), was overlaid onto all experimental conditions. The relationship between the time shift (a_T) and temperature (T) was found to be linear, with an R^2 value of 0.9502 (Eq. 3.10). The final generalized Midili Kucuk equation for steamed drying, as presented in Eq. 4.1, was utilized to estimate the moisture ratio at 50, 60, and 40 °C. A comparison of the measured and (Fig. 4.1.3a, 4.1.3b, 4.1.3c) showed that slight deviation of the predicted curve of 40 °C while later, the measured values and the predicted curve coincided adequately for 60 °C.

$$MR = (1.009e^{-0.067t^{0.782}} + 4.54 \times 10^{-4}t)(-0.3002 + 0.025t) \quad (4.1)$$

Table 4.1.2: Midili-Kucuk coefficient at different temperatures

Temp	40 °C	50 °C	60 °C
R^2	0.990	0.997	0.991
RMSE	0.238	0.120	0.230
a	1.006	1.009	1.008
b	9.94×10^{-5}	4.54×10^{-4}	4.83×10^{-4}
k	0.053	0.067	0.087
n	0.759	0.782	0.729

4.1.1.3 Diffusion equation

This diffusion equation studied the migration of water (diffusant A) into the solid system (rice). The volumetric concentration with respect to time showed rapid uptake of water for the first 1 h and then a slow uptake until it reached saturation after 3 h (Fig. 4.1.4). The grain parameters along with their measurement method that were considered for solving were tabulated in Table 4.5. The diffusion model (Eq. 3.9) was solved using the discretized solution of the finite difference method.

The predicted fractional water absorption estimated using the solved diffusion equation almost coincided with the experimental value (Fig. 4.1.5). The sum of squared errors (SSE) between the predicted and experimental was found to be 1×10^{-6} , which was considerably low suggesting us a good optimization of parameters. The effective diffusivity coefficient was found to be $5.43 \times 10^{-11} \text{ m}^2/\text{s}$ for the initial concentration (Fig. 4.1.6), and it kept on increasing with an increase in concentration. The moisture progressed radially through the surface, and it required 1.5 h to reach the core from the surface when the moisture value had reached approximately 30% (wb) (Fig. 4.1.7). After 2 hours, the soaked grain tended to reach saturation. The saturation moisture content was calculated using two methods: Peleg's method and Bello's method [5].

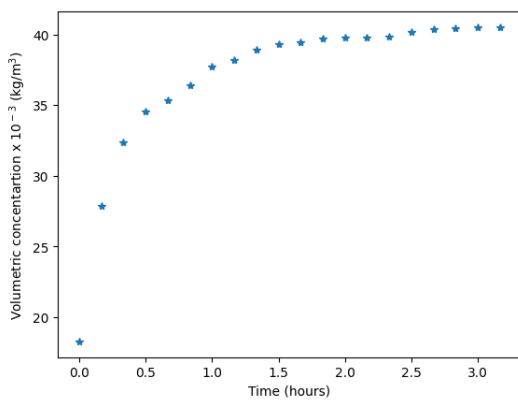


Fig. 4.1.4: Change in volumetric concentration during soaking

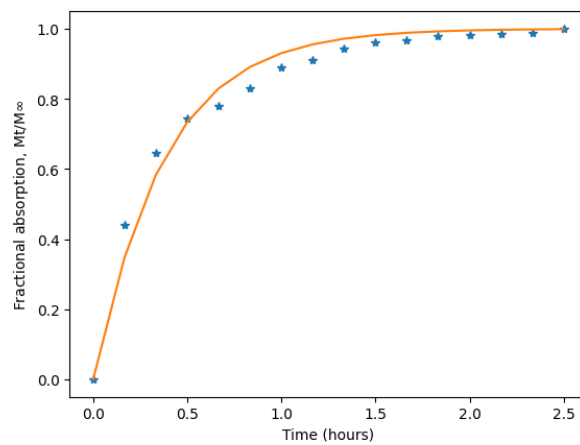


Fig. 4.1.5: Fractional absorption of moisture with respect to time for soaking at 60 °C

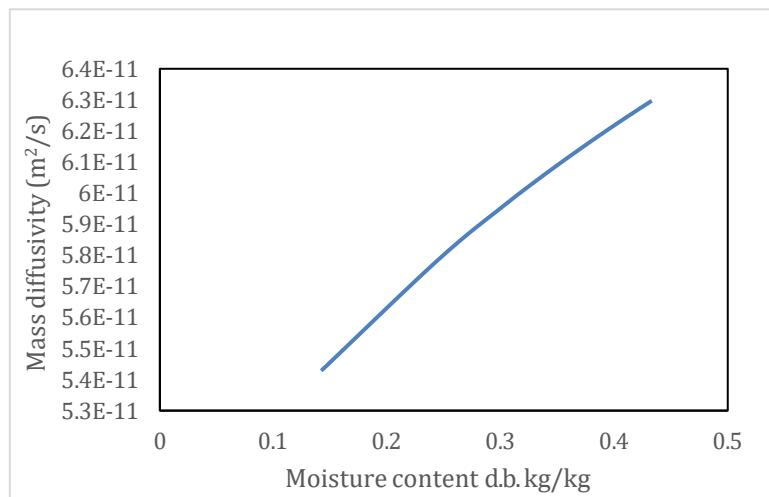


Fig. 4.1.6: Mass diffusivity vs moisture content at 60 °C

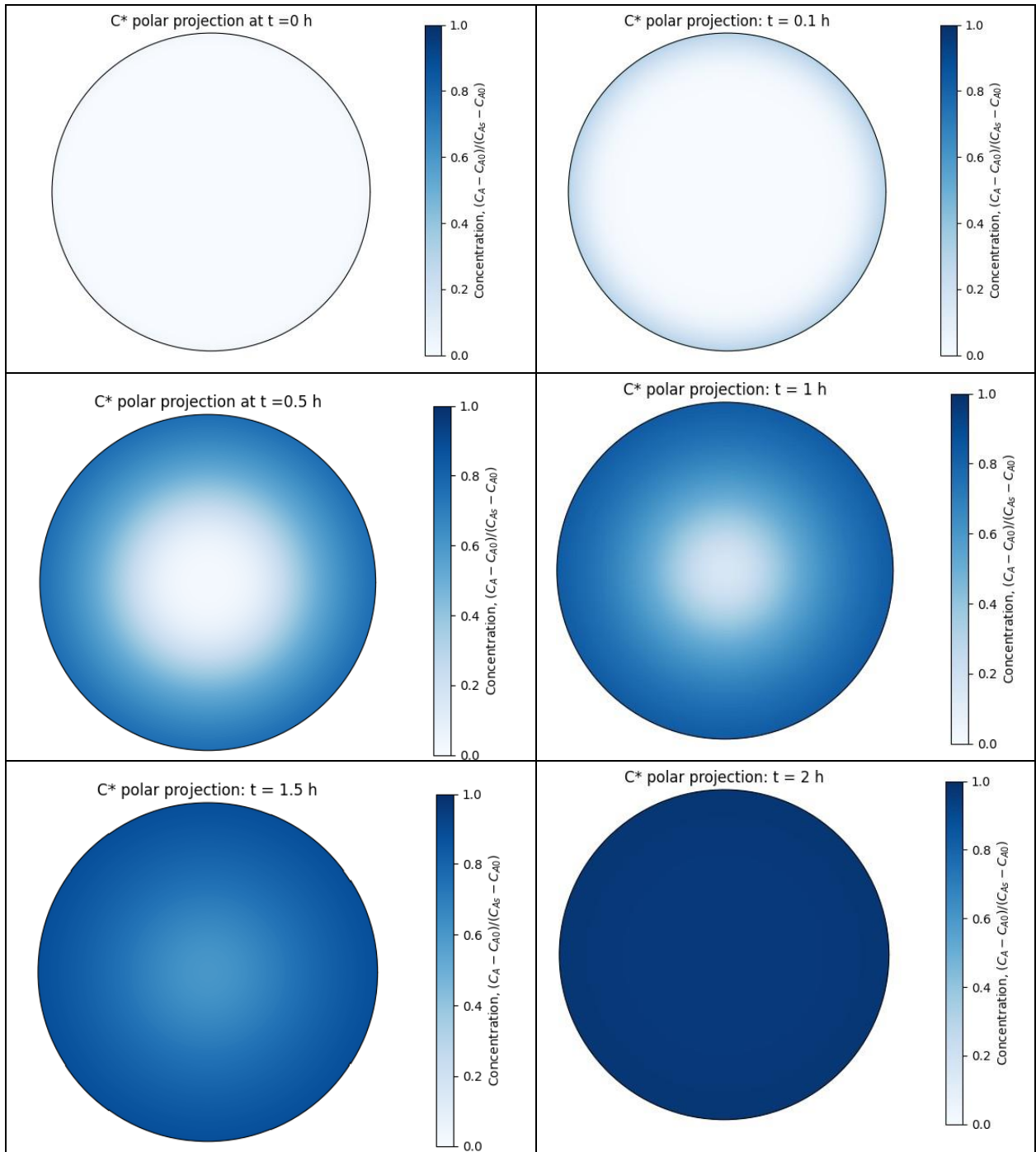


Fig. 4.1.7: Color mapping of concentration of water radially along the polar coordinates from surface to core at 60 °C

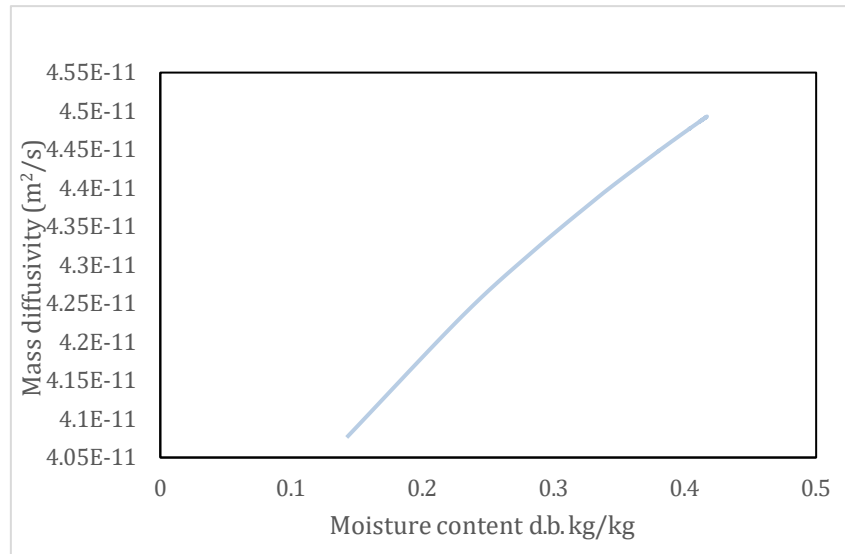
The physical parameters that were measured are listed in Table 4.1.3. Effective diffusivity independent of concentration (D_0) was calculated accordingly for each temperature (Table 4.1.4). The change in diffusivity coefficient for soaking at 40 and 50 °C was also calculated. The diffusivity range with a change in moisture content for 50 °C was $4.05\text{-}4.50 \times 10^{-11} \text{ m}^2/\text{s}$ (Fig. 4.1.8). And the diffusivity range with a change in moisture content for 40 °C was $3.30\text{-}3.60 \times 10^{-11} \text{ m}^2/\text{s}$ (Fig. 4.1.9). Appendix II shows the changes in moisture concentration radially at 40 and 50 °C.

Table 4.1.3: Grain dimensions and physical parameters

Physical parameters	Value	Measurement device
Length of rice grain(m)	6.02×10^{-3}	Grain scanner
Width of rice grain (m)	2.64×10^{-3}	Grain scanner
Thickness of rice grain (m)	1.69×10^{-3}	Vernier Caliper
Density of solid in rice grain (kg/m^3)	1572	Gas Pycnometer
Density of water (kg/m^3)	1000	Considered temperature independent [5]

Table 4.1.4: Saturation moisture and Diffusivity comparison

Temperature (°C)	Saturation moisture content (Peleg method)	Diffusivity D independent of concentration (D_0) (m^2/s)	Saturation moisture content (Bello Method)	Diffusivity D independent of concentration (D_0) (m^2/s)
40	0.4408	5.32×10^{-11}	0.3209	2.83×10^{-11}
50	0.4508	6.64×10^{-11}	0.4039	3.73×10^{-11}
60	0.4613	7.92×10^{-11}	0.4136	4.73×10^{-11}

**Fig. 4.1.8: Mass diffusivity vs moisture content at 50 °C**

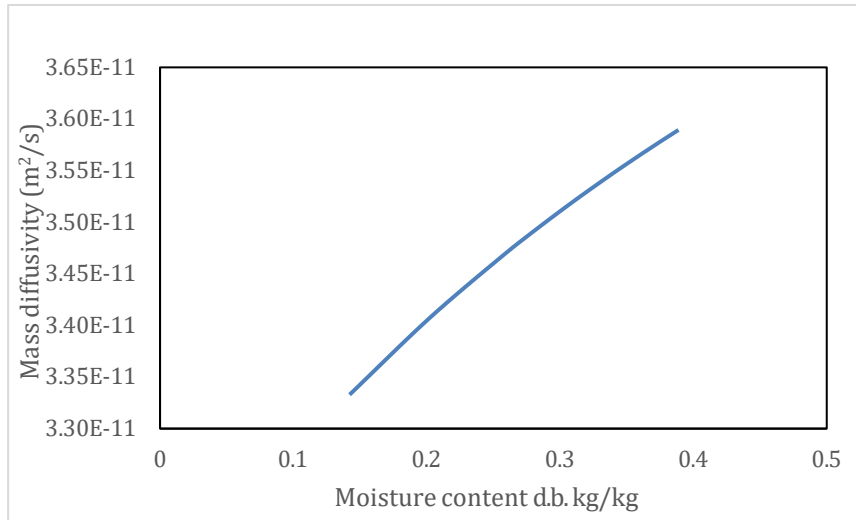


Fig. 4.1.9: Mass diffusivity vs moisture content at 40 °C

4.1.2 Steaming Kinetics

The degree of gelatinization percentages for each treatment pressure (0, 0.05, 0.1, 0.15, and 0.2 MPa gauge pressure) were plotted against the treatment times (Fig. 4.1.10).

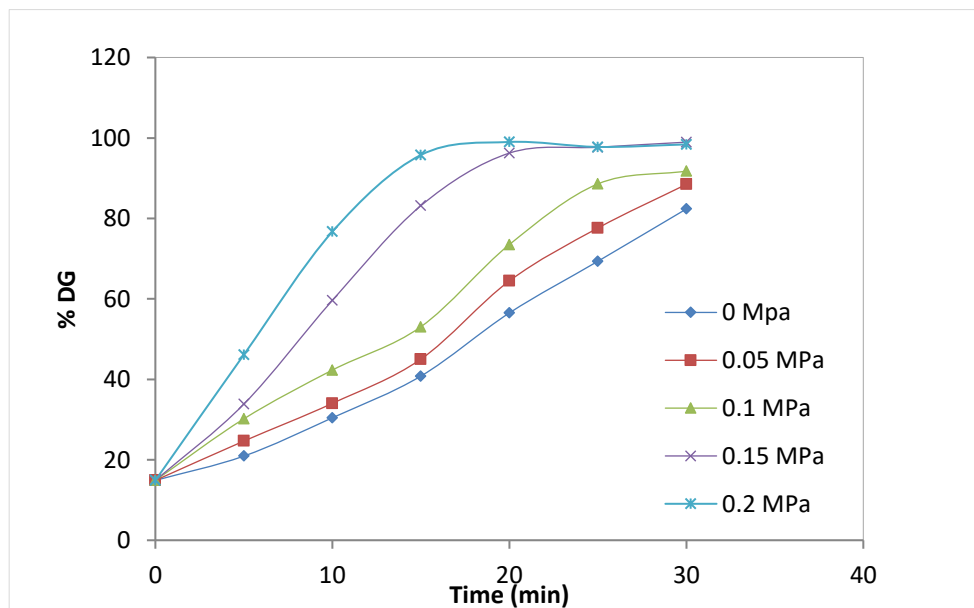


Fig. 4.1.10: Rate of gelatinization at different pressures for Komal Chaul at different treatment time

The curves in Fig. 4.1.10 suggest that DG values increased exponentially with time of treatment and moreover rate of gelatinization was dependent on steam pressure. Initially, before subjecting it to pressure treatment (i.e., 0 min), the degree of gelatinization was found to be

13.29% for the soaked rice, therefore, suggesting that some degree of gelatinization had occurred before steaming while soaking at 60 °C for 90 min. It had also been observed that with an increase in pressure, the gelatinization rate increases thus requiring a lesser time to achieve complete gelatinization. For instance, at 0 MPa to achieve more than 50% gelatinization, it took around 12 min, whereas at 0.2 MPa within 9 min, the DG % obtained is more than 95%. Nearly, a similar result has been found for rough rice parboiling of *Komal Chaul* at 1 atm (approx. 0.1 MPa) by Wahengbam, 2020, which is 99% gelatinization at 20 min[6].

4.1.2.1 Gelatinization reaction kinetics

The percentage of DG change was expressed as a function of time with a reaction rate constant (Eq. 3.35) for determination and expressed as a logarithmic expression by integrating Eq. 3.36. The reaction rate constant was obtained by finding the slope of Eq. 3.37 (Fig. 4.1.11).

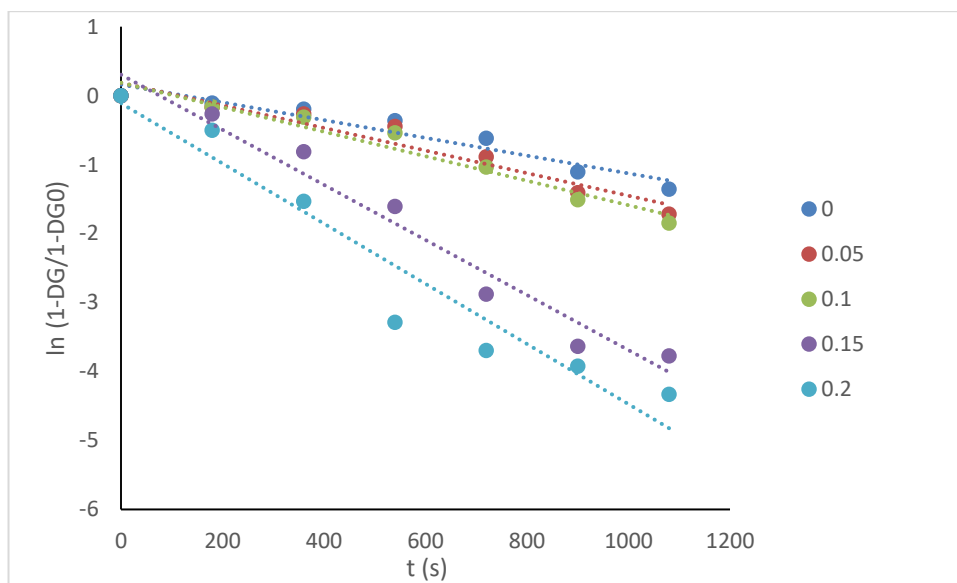


Fig. 4.1.11: Rate constant estimation plot

The reaction rate was obtained for each pressure treatment by fitting the experimental data in Eq. 3.37. The calculated reaction rate constant (k) and the performance of the fitted model based on R^2 are presented in Table 4.2.1. It can be observed that a goodness of fit was obtained (Fig. 4.1.12) since the R^2 values were much closer to 1 and RMSE values were low for all the data. Earlier reports also suggested that the rice starch gelatinization process follows the first-order reaction very well [7,8]. The k values had been observed to vary between 0.003 to 0.001 s^{-1} with varying pressure. The maximum changes were due to the increase in steaming time. With increasing pressure, the reaction rate constant shows a tendency to increase, which

had been confirmed by other studies [9,10]. The activation energy for the steaming process was found to be 18.7×10^2 kJ/mol.

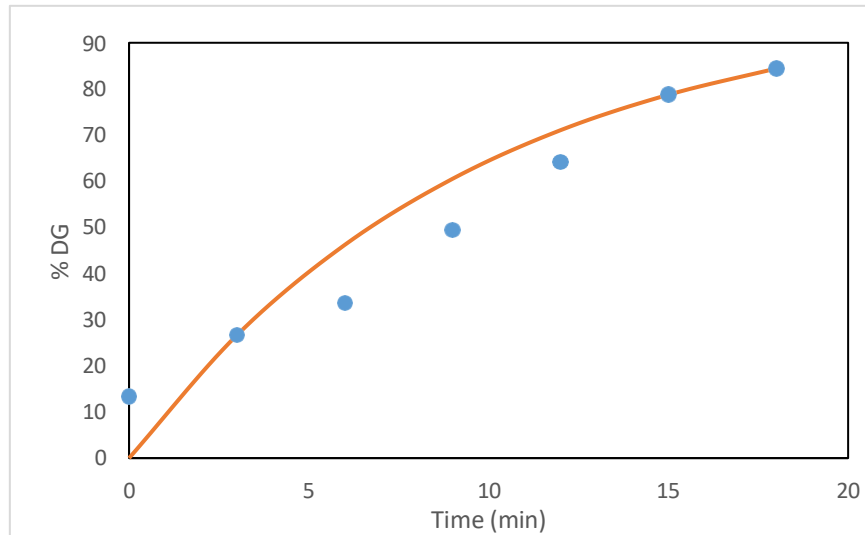


Fig. 4.1.12: Degree of gelatinization against time for measured and predicted at 0.15 MPa

Table 4.2.1: First-order kinetics rate equation accuracy metrics

Press (MPa)	k (s^{-1}) $\times 10^{-3}$	R^2	RMSE (fraction)
0	1.288	0.927	1.33
0.05	1.637	0.936	0.88
0.1	1.776	0.955	0.99
0.15	3.994	0.964	2.4
0.2	4.363	0.929	4.6

Also, a multilinear regression was performed to know the effect of processing variables on the percentage of DG values. The multilinear regression model showed significance with a p-value < 0.01 and quite a decent accuracy with an R^2 value of 0.994, adjusted R^2 value of 0.884, and SSE value of 10.22 %. The cause of change in DG values with respect to pressure and temperature can be expressed in the form of a linear equation (Eq. 4.2) with the coefficients obtained as:

$$DG (\%) = 0.9 + 171time + 4.4pressure \quad (4.2)$$

This gives a clear idea that steaming time tremendously affects the gelatinization process as compared to steaming pressure. The Eq. 4.3 implies an elevated moisture content with an increase in steaming time. When the temperature favors the gelatinization process, the

progression of the gelatinization process depends on the availability of water molecules at the reaction site. Accordingly, the measured data on DG and moisture content leads to a linear relationship, for the holding time for the steaming process as given in Eq. 4.3.

$$DG = 5.3615M.C. - 156.22 \quad (4.3)$$

Where DG is the degree of gelatinization percentage and M.C. percentage on a wet basis.

4.1.2.2 Steam diffusion kinetics

The steam diffusion kinetics at different pressures were simulated using Fick's law equation (Eq. 3.16) based on the initial and boundary conditions as mentioned earlier in section 2.2. Considering the volume change to be negligible, the diffusion of steam was in the radial direction with constant diffusivity. In order to obtain a solution for the PDE (Eq. 3.16) the grain dimensions were considered to be in spherical coordinates. The equivalent diameter calculated for the rice grain prior to steaming was 0.0016 m. The initial moisture content obtained for soaked rice was 43.09% (db). The average moisture content at different time intervals for each pressure treatment as shown in Fig. 4.1.13 gives an idea about the rate of steam penetration in the rice grain. It had been determined that with the rise in time, the moisture content increased exponentially; and with a rise in pressure, the steam penetration additionally showed an increase till it reached saturation. This can be explained in reference to Miah et al. [11] which describes that higher pressure allows a greater amount of steam and creates more kinetic energy for penetration inside the grain, thus increasing the gelatinization front.

In order to estimate the diffusivity coefficient, the solution for Fick's equation i.e., Eq. 3.37 was used, where the average moisture content was a function of equilibrium moisture content and time. The reaction rate constant (k) values were obtained from Eq. 3.29, the specific heats were calculated using Eq. 3.40, 3.41, and 3.43, and the saturated temperatures for steam for every treatment pressure were obtained using the steam table. The equilibrium moisture contents were calculated using the relation suggested by [7] as in Eq. 3.44, and the values obtained at respective pressure (0, 0.05, 0.1, 0.15, 0.2 MPa) were 0.394, 0.511, 0.537, 0.555, and 0.567. Further, these calculated values were used in Eq. 3.45 to estimate the diffusivity values, and the D_{eff} values were calculated using the SciPy library in Jupyter Notebook to solve the nonlinear series equation. The D_{eff} values obtained at 0, 0.05, 0.1, 0.15 and 0.2 MPa were 1.76, 3.52, 5.28, 7.04, and $8.8 \times 10^{-8} \text{ m}^2/\text{s}$ respectively. Hence, there was a rise in the diffusivity of steam with the rise of steam pressure; and this statement can be validated by an earlier report from Miah et al. [11]. As suggested by another study [12], essentially, the diffusion coefficient is a function of the

absolute temperature, increased by the power of 1.5 pressure inside the grain.

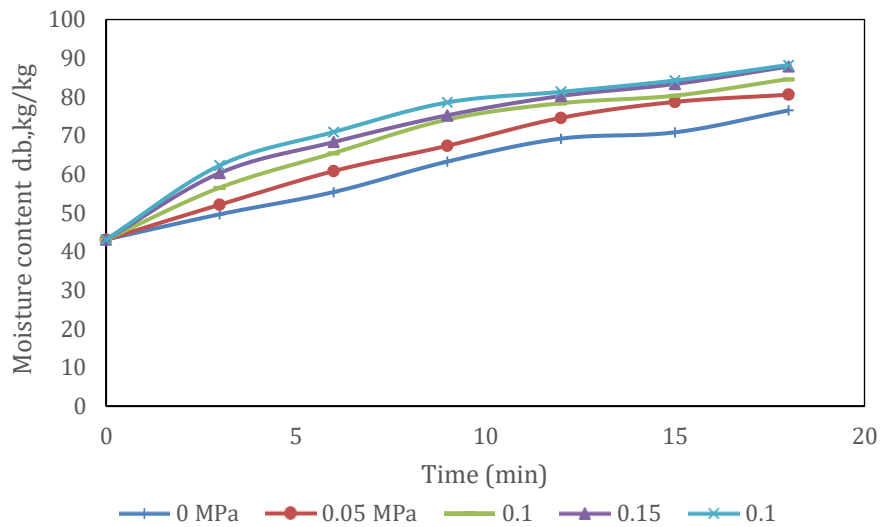


Fig. 4.1.13: Moisture content at different pressures of steaming

4.1.3 Drying kinetics

The rate of moisture removal at three different temperatures was studied. The initial moisture content of the steamed brown rice was around 37.21% (wb) (59.27% db). The equilibrium moisture content for drying was measured to be 9.31% (wb) (10.27% db), 10.28% (wb) (11.24% db), and 11.01% (wb) (12.37% db) at 60 °C, 50 °C, and 40 °C. The change in moisture content with respect to time is shown in Fig. 4.1.14.

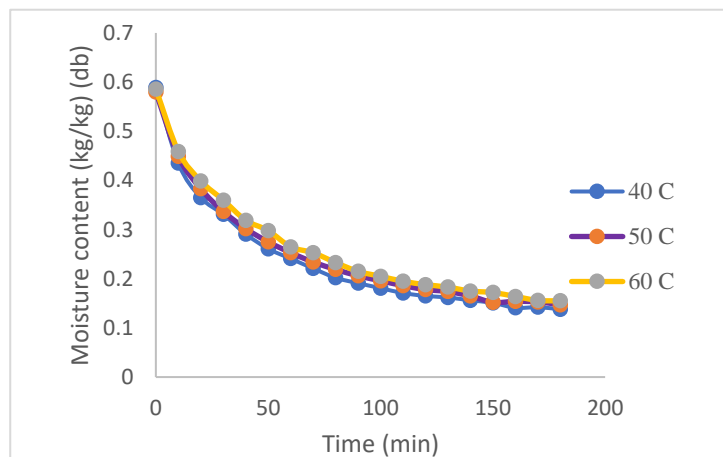


Fig. 4.1.14 Change in moisture content w.r.t. time (a) at 60, 50 and 40 °C during drying of steamed brown rice.

The moisture content abruptly decreases during the initial stages and later tends towards saturation while drying.

4.1.3.1 Empirical drying models

Usually for food drying, considering the Biot number to be less than 0.1, the ambient air temperature and rice temperature were the same. All the models referenced in Table No. 3.2 were tested against the drying data at 60, 50, and 40 °C. Among these, the Page model outperformed others with an R^2 value greater than 0.99 and a RMSE less than 0.01. The performances of other models of 60 °C are shown in table No. 4.3.1. Similarly, for drying data of 50 °C and 40 °C Page equation fitted the best, with an R^2 value greater than 0.99 and an RMSE less than 0.01 were obtained as detailed in Table 4.3.2 and 4.3.3 respectively. Using the predicted equation, the predicted moisture content for each time was calculated using the developed Page model. The predicted moisture content was plotted against time and compared with the experimental data; it could be observed that they coincide very well (Fig 4.1.15). In most cases of intermittent drying of rice, the Page model fits the best [13]. The time predicted to reach 13% m.c. (wb) was found to be 150, 165, and 185 min at 60, 50, and 40 °C respectively.

Table 4.3.1: Showing the performances of the models for 60°C data.

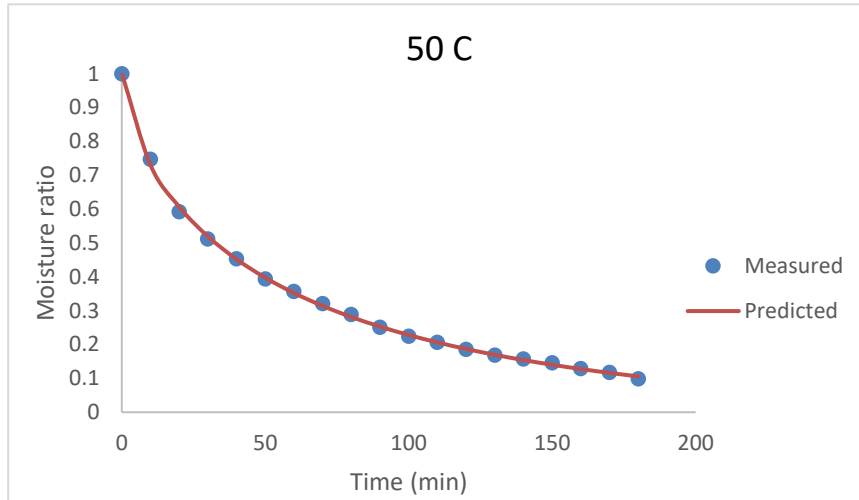
Model names	Model coefficients and constants	R^2	RMSE (fraction)
Page	k=0.0782 n=0.6805	0.999	0.006
Newton	k=0.01644	0.939	0.075
Modified Page	k=0.06554 n=0.7292	0.938	0.014
Henderson and Pabis	k=0.01572 a=0.9587	0.971	0.051
Wang and Singh	a=-0.0124 b=4.126×10 ⁻⁵	0.887	0.103
Diffusion	k=0.01082 a=0.07818 b=1.396	0.887	0.102

Table 4.3.2: Showing the performances of the models for 50°C data.

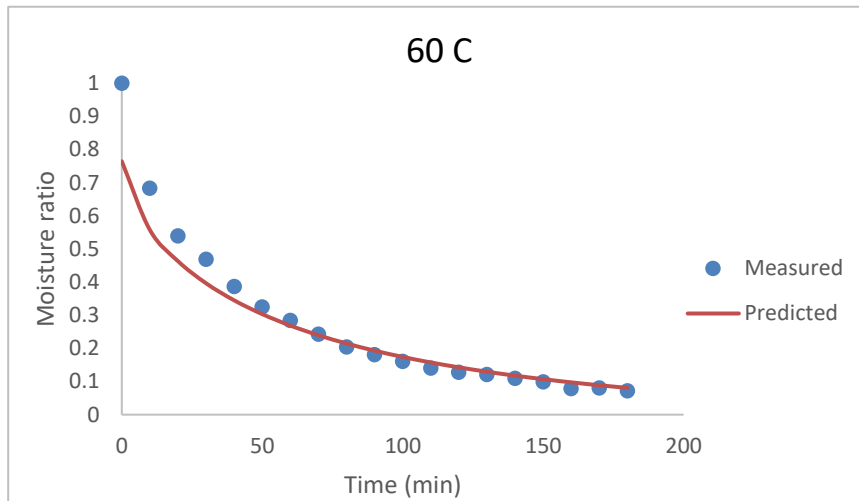
Model names	Model coefficients and constants	R²	RMSE (fraction)
Page	k=0.0643 n=0.715	0.998	0.007
Newton	k=0.017	0.978	0.037
Modified Page	k=0.06468 n=0.6802	0.988	0.028
Henderson and Pabis	k=0.01699 a=0.9995	0.987	0.029
Wang and Singh	a=-0.01277 b=4.332×10 ⁻⁵	0.797	0.115
Diffusion	k=0.01677 a=0.5336 b=1.029	0.796	0.115

Table 4.3.3: Showing the performances of the models for 40°C data.

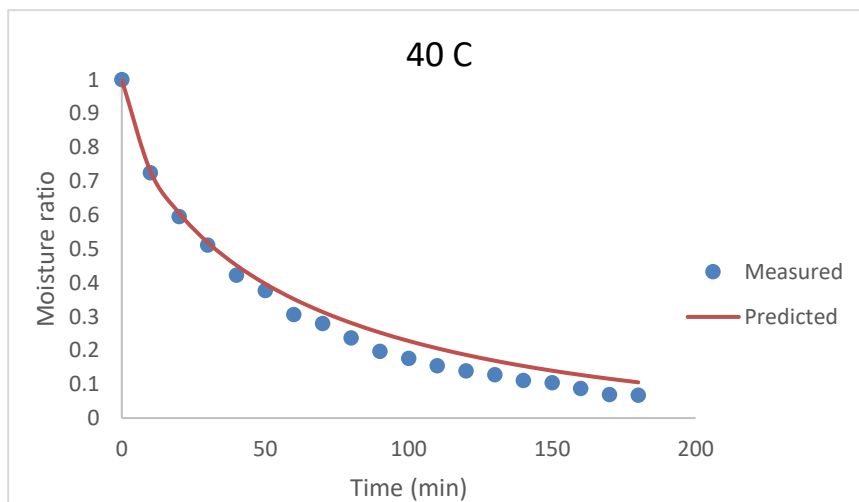
Model names	Model coefficients and constants	R²	RMSE (Fraction)
Page	k=0.0540 n=0.7502	0.991	0.009
Newton	k=0.02017	0.981	0.030
Modified Page	k=0.0438 n=0.7043	0.9851	0.027
Henderson and Pabis	k=0.0211 a=1.045	0.989	0.024
Wang and Singh	a=-0.01412 b=5.021×10 ⁻⁵	0.909	0.067
Diffusion	k=0.01468 a=3.221 b=0.872	0.909	0.067



a) 50°C



a) 60°C



c) 40°C

Fig. 4.1.15: Predicted values of moisture content Page equation and comparison with measured values

4.1.3.2 Generalized drying equation for time temperature relation

The Midili Kucuk equation consists of two terms; one is the linear function of time, and the other one is an exponential term (Eq. 3.26). The coefficients for the Midili Kucuk equation for the temperature 40, 50, and 60 °C were found using curve fitting. The fit of the drying moisture ratio at all temperatures was good with R^2 greater than 0.99. The RMSE values of the predicted model showed a better prediction accuracy (Table 4.3.4).

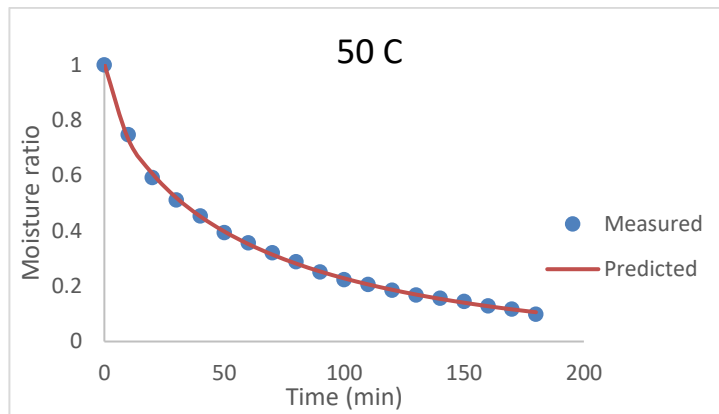
Table 4.3.4: Statistical metrics of Midili-Kucuk prediction

Temperature	R^2	RMSE (fraction)
40 °C	0.998	0.009
50 °C	0.999	0.004
60 °C	0.998	0.007

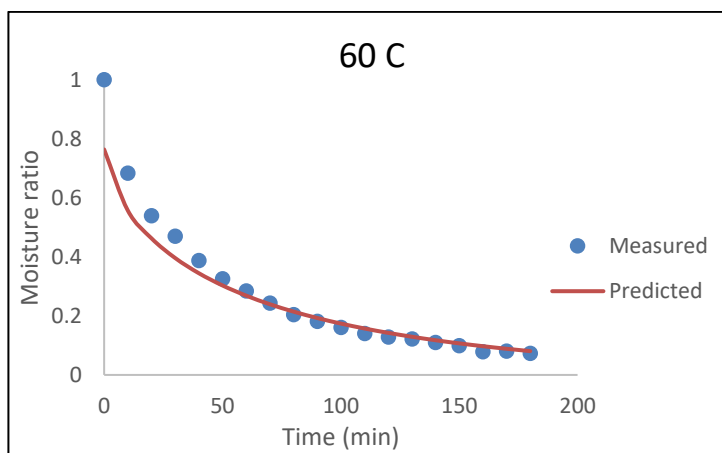
Table 4.3.5: Midili-Kucuk coefficient at different temperature

Temperature →	40 °C	50 °C	60 °C
Coefficients ↓			
a	1.006	1.002	0.999
b	-9.94×10^{-5}	-4.54×10^{-4}	-4.83×10^{-4}
k	0.053	0.068	0.082
n	0.759	0.663	0.658

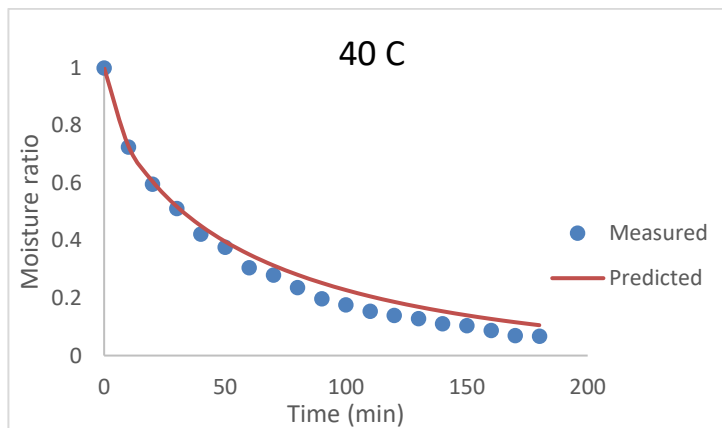
The values of exponential coefficient (a), rate constant (k), power factor (n), and linear constant were obtained as tabulated in Table 4.3.5. The developed model was shifted to a reference equation as suggested by Bezbaruah and Hazarika [4]. The rate constant (k) was found to increase with an increase in temperature, while the power factor of time (n) showed a decreasing trend. This is because the drying rate constant is an exponential function of temperature [4]. The master curve i.e., the 50 °C (central condition) predicted equation was superimposed on all experimental conditions from the calculated time shift (using Eq. 3.27) to a central condition. The relationship between the time shift (a_T) and temperature T was found to be a linear equation with an R^2 value of 0.9502 (Eq. 3.27). The final generalized developed Midili Kucuk equation for drying of steamed brown rice represented in Eq. 4.2 was used to estimate the moisture ratio at 50, 60, and 40 °C and the comparison plots between the measured and predicted (Fig. 4.1.16a, 4.1.16b, 4.1.16c) showed that slight deviation of the predicted curve at the initial measurement while at later time, the measured values and the predicted curve coincides adequately.



(c) 50°C



(d) 60°C



(c) 40°C

Fig. 4.1.16: Predicted values of moisture using developed Midili Kucuk equation

An earlier study on the same rice showed that feed-forward artificial neural networks showed well accuracy in relating the moisture ratio as a function of temperature and time [14]. The final form of the equation is given as Eq 4.1[#]

$$MR = (1.002e^{-0.067t^{0.782}} + 4.54 \times 10^{-4}t)a_T$$

$$\text{Where } a_T = 3.0709 - 0.0393T \quad (4.1^{\#})$$

4.1.3.3 Diffusion equation for drying kinetics

The diffusion model is based on the phenomenon of mass transfer that studies the movement of moisture during the process of absorption or desorption of moisture. The diffusion equation was solved using a finite difference method for each temperature (40, 50, and 60 °C). Firstly, the Cranks solution was used where the constant moisture diffusivity coefficients at 40, 50, and 60 °C were found to be 7.77×10^{-9} , 5.2×10^{-9} , and 2.6×10^{-9} m²/s, respectively. Using the Arrhenius equation, the activation energy was estimated to be 24.9×10^{-2} kJ/mol. Hsu method was applied to solve the partial differential method (Eq. 3.18 - Eq. 3.24). The diffusion model using the Hsu model predicted well. For example, at 60 °C the fractional moisture values of the prediction curve and experimental points coincided very well (Fig. 4.1.17). The prediction and experimental comparison of fractional moisture values at 40 and 50 °C were illustrated in Appendix V. The pre -exponential factor diffusivity (D'_0) was found to be 1.62×10^{-7} , 2.24×10^{-6} , and 3.32×10^{-6} at 40, 50, and 60 °C. D_{eff} with respect to time were in the range of 2.72×10^{-10} to 3.40×10^{-9} m²/s with increasing temperature. The change in moisture diffusivity values with respect to time at different temperatures are given in Appendix VI. Wahengbam et al., [6] found diffusivity in the range of 2.08×10^{-10} to 3.34×10^{-10} m²/s with increasing temperature. The reason for the slight variation could be attributed to the efficiency of the drying system. The diffusivity was higher at the initial time and then slowed down as time increased suggesting saturation. Azzouz and co-workers [15] also observed the temperature dependencies of the diffusivity coefficient while drying grapes. At the initial stage of drying, the mass diffusion from the rice grain was rapid and later remained stable, which is due to the air temperature attaining the set value of temperature [16]. Therefore, with the approach of the finite difference method and particle swarm optimization, the partial differential equation for mass diffusion was solved, which helped in estimating the moisture content throughout the grain at different time intervals.

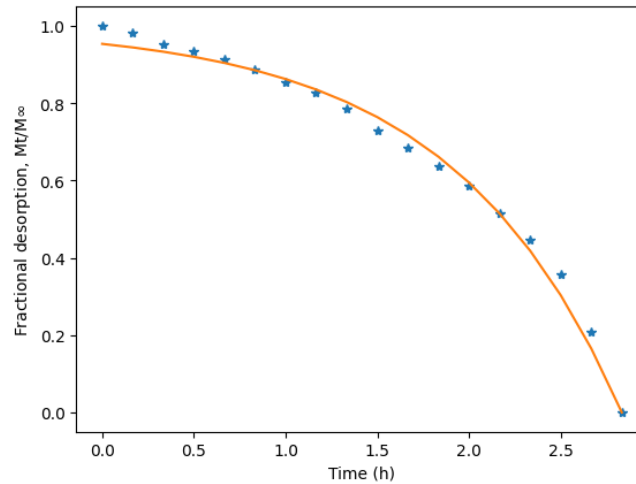


Fig. 4.1.17: Fractional desorption of moisture with respect to time for drying at 60 °C

4.1.4 Rehydration kinetics

The rehydration kinetics was done at 40, 50, and 60 °C for the time duration of 25 min and moisture was measured at 2.5 min intervals. The saturation moisture content obtained during the process of rehydration at 40, 50, and 60 °C were 0.541, 0.690, 0.789 kg/kgdb reached at time 15 min, 22.5 min and 25 min respectively. The average effective diffusivity was 3.39×10^{-6} , 2.24×10^{-6} , and 1.62×10^{-6} m²/s at 60, 50, and 40 °C respectively. The activation energy for the process was estimated to be 78.48 kJ/mol. It ensures that the water absorption or hydration of *Komal Chaul* was much higher than *Chokuwa* rice hydration attributing to the ready to the property of *Komal Chaul*.

4.1.5 Calibration of processing parameters using spectra

4.1.5.1 Moisture content calibration for soaking process

The calibration process involved finding integrated absorbance values from spectra at different concentrations. Average reflectance values of spectra over the wavelength range 740-1050 nm were chosen for 60, 50, and 40 °C for 0-120 min at 20 min intervals for moisture values ranging from 13-33 % (wb). The plot for average absorbance values against wavelength obtained at different times for soaking for 60 °C as shown in Fig. 4.1.18, gave the idea about the varying reflectance values and the peak range 925-1025 nm corresponding to water molecule [17]. The reflectance values were converted into absorbance values for the time considered and the integrated absorbance for each concentration was obtained using Origin Pro 8.5. The integrated absorbance values with respect to time were plotted and found to follow an almost linear trend (Fig. 4.1.19).

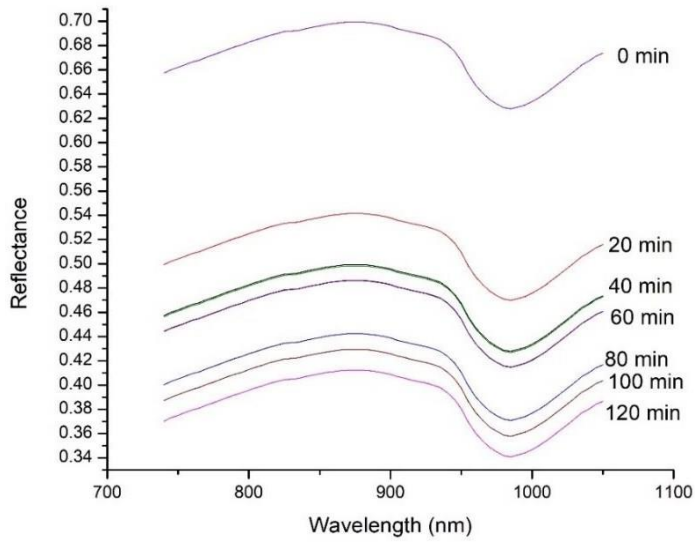


Fig. 4.1.18: Reflectance values against wavelength for soaking process at 60 °C

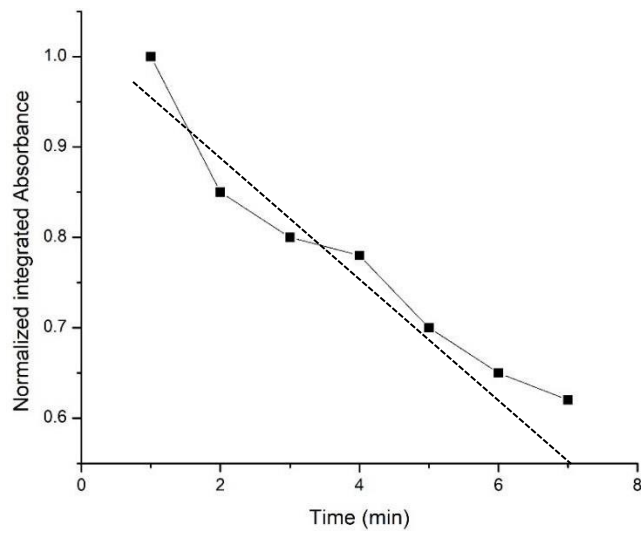


Fig. 4.1.19: Normalized integrated absorbance vs time for soaking spectra at 60 °C

Table 4.3.6: Coefficients and statistical metrics of linear plots of soaking spectral calibration

Temperature	Coefficients	R ²	RMSE (%)
40 °C	a=0.428, b=-0.451	0.910	0.0014
50 °C	a=0.398, b=-0.524	0.932	0.0011
60 °C	a=0.383, b=-0.621	0.926	0.0013

The normalized integrated absorbance values were then linear fitted with the fractional concentration of moisture content (wb). The slope (b) and intercept (a) of the linear equations were obtained for 40, 50, and 60 °C as well, and goodness of fit was expressed in terms of R^2 and RMSE (Table 4.3.6). Slope and intercept were found to decrease with an increase in temperature. This may be because in the case of soaking the reflectance values were found to decrease with an increase in time.

4.1.5.2 Degree of gelatinization calibration for the steaming process

A similar process of calibration for the degree of gelatinization was followed. The integrated absorbance values were calculated by integrating the area of the peak contributing to water molecule in the range of 925-1025 nm at 0, 3, 6, 9, 12, and 15 min of steaming (Fig. 4.1.20). The normalized value of integrated absorbance and degree of gelatinization values at respective pressures of 0 MPa showed a linear trend (Fig. 4.1.21), and similarly at 0.05, 0.1, 0.15 and 0.2 MPa gauge pressure the integrated absorbance values were fitted linearly against concentration of the respective time. The coefficients and statistical metrics are represented in Table 4.3.7. Similar to soaking, steaming spectra showed a decreasing trend with increasing time while the peaks were sharper in the case of steaming.

4.1.5.3 Moisture content calibration for drying process

The drying process was calibrated using a laboratory-developed sensor discussed in Appendix-I. The NIR zone of this sensor is from 700-940 nm. Reflectance peaks were observed between 800-900 nm. Therefore, integrated absorbances were calculated by integrating the area of absorbance values from 800-900 nm (Fig. 4.1.22). The normalized integrated absorbance was calculated by integrating the area of absorbances against wavelength (Fig. 4.1.23) and they were linearly correlated with moisture values obtained at drying time. The coefficients and statistical metrics as mentioned in Table 4.3.8. In the case of drying the peaks showed an increasing trend with time. The increase in absorbance values with time for respective concentrations showed an opposite trend as soaking and steaming. This may be due to the desorption of water from the rice grains during drying.

The conventional old chemometric technique considers discrete wavelength specifying a single bond, whereas the other bonds may also contribute to the attribute of concern. Therefore, our next objective is multidimensional calibration using machine learning techniques.

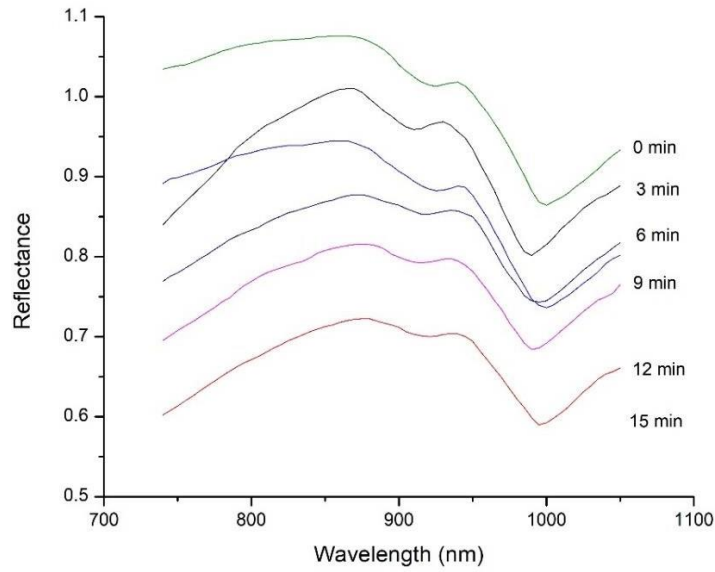


Fig. 4.1.20: Reflectance values against wavelength for steaming process at 0 MPa

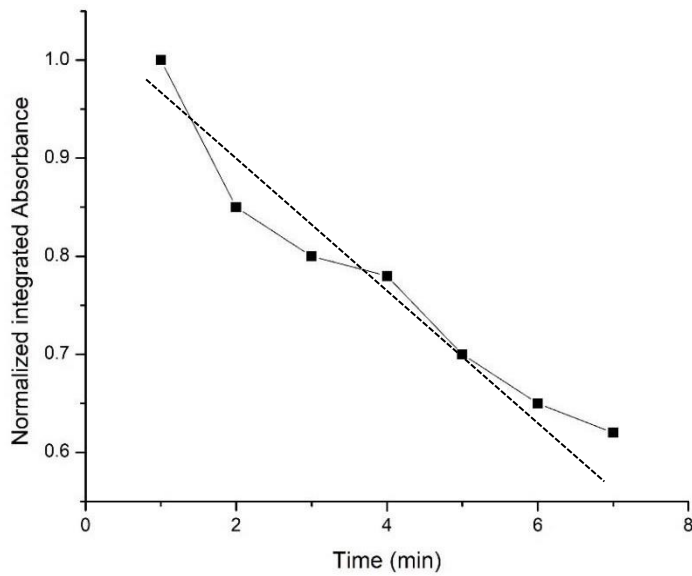


Fig. 4.1.21: Normalized integrated absorbance vs time for steaming spectra at 0 MPa

Table 4.3.7: Coefficients and statistical metrics of linear plots of steaming spectral calibration

Pressure	Coefficients	R ²	RMSE (%)
0	a=0.510, b=-0.695	0.991	0.0012
0.05	a= 0.492, b=-0.785	0.993	0.0011
0.1	a=0.438, b=-0.803	0.995	0.0013
0.15	a=0.421, b=-1.41	0.894	0.0101
0.2	a=0.398, b=-1.49	0.892	0.0110

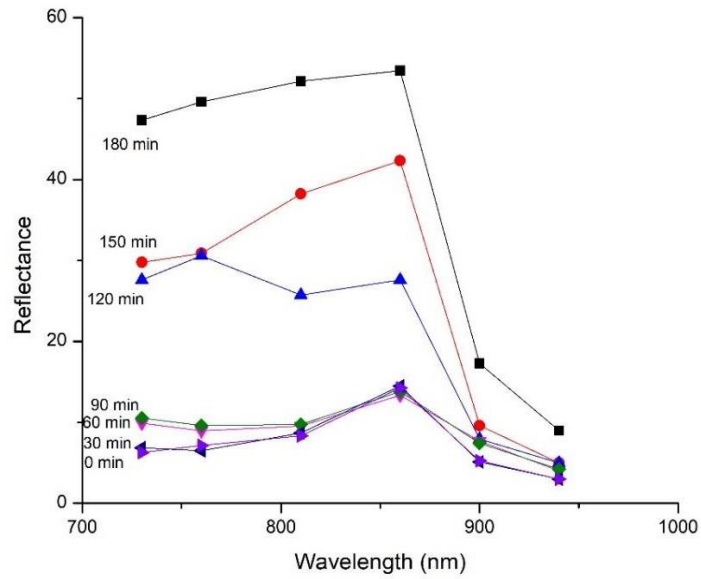


Fig. 4.1.22: Reflectance values against wavelength for drying process at 60 °C

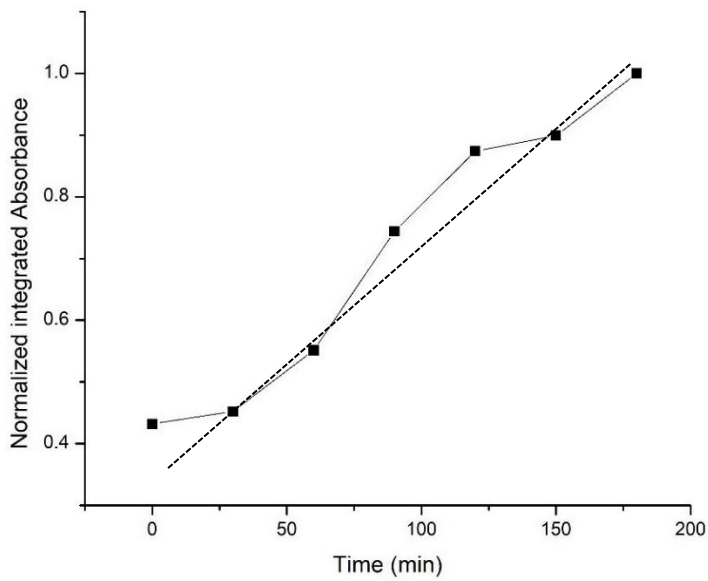


Fig. 4.1.23: Reflectance values against wavelength for drying process at 60 °C

Table 4.3.8: Coefficients and statistical metrics of linear plots of drying spectral calibration

Temperature	Coefficients	R ²	RMSE(%)
40 °C	a=0.528, b=1.451	0.951	0.0014
50 °C	a=0.698, b=1.628	0.962	0.0011
60 °C	a=0.723, b=1.821	0.946	0.0013

4.2 Results and discussion for spectral calibration of parameters of the parboiling process using ML models

4.2.1 Soaking spectral modeling

At every temperature i.e., 40, 50, and 60 °C, and at every time interval of 10 min throughout the 0–180-minute soaking process, each spectrum was obtained in triplicates giving us a total scan of 171 (3×19×3). The spectra were acquired using the software development kit SCiO Lab. The moisture content at each time for every temperature was measured. The raw reflectance values of soaked rice at every condition against wavelength were plotted (Fig. 4.2.1). PLS was used as a regression model for estimating the quantitative value of moisture during soaking of rice. The machine learning model was trained using the training data set. K-fold validation was used since the dataset consists of more than 100 spectra. The training set at k=1, k=2, and k=3 was validated in reference to the number of latent variables. It was observed that the RMSE values does not change after the 5 latent variables in Fig. 4.2.2. The regression model of PLS didn't fit well for NIR spectral data with statistical metrics of performance R^2 :0.305, RMSE: 2.109 % for the validation dataset and R^2 : 0.338, SEP: 2.109 for the test dataset. The low R^2 and low RMSE of prediction values suggested very little correlation between the variables and low error. The green line plot, which is the expected zero error relation showed the regressive relation between the measured and the predicted represented by the blue plot was hardly correlating (Fig. 4.2.3).

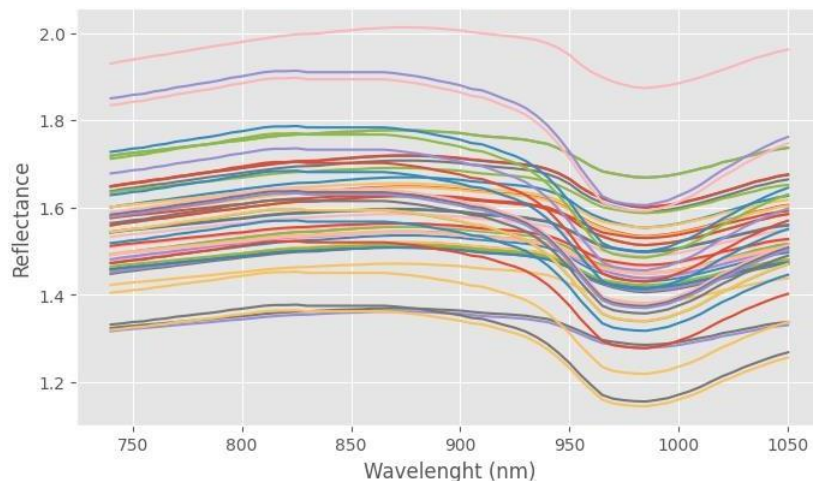


Fig. 4.2.1: Raw reflectance against wavelength during soaking

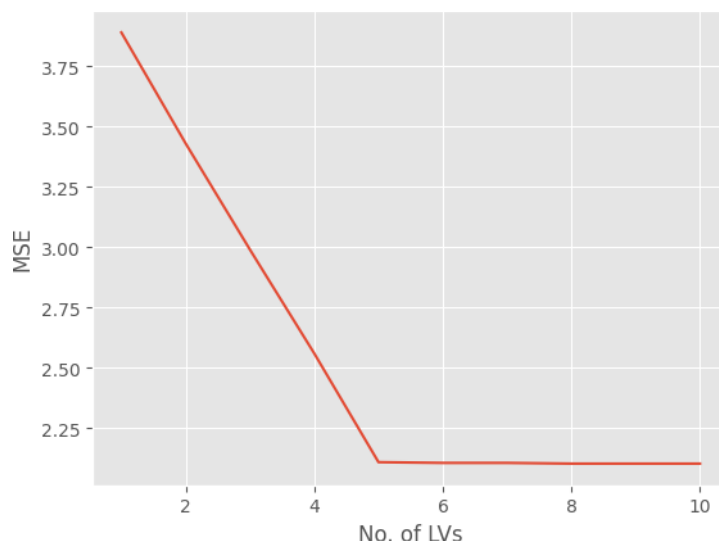


Fig. 4.2.2: Validation and hyper tuning of LVs

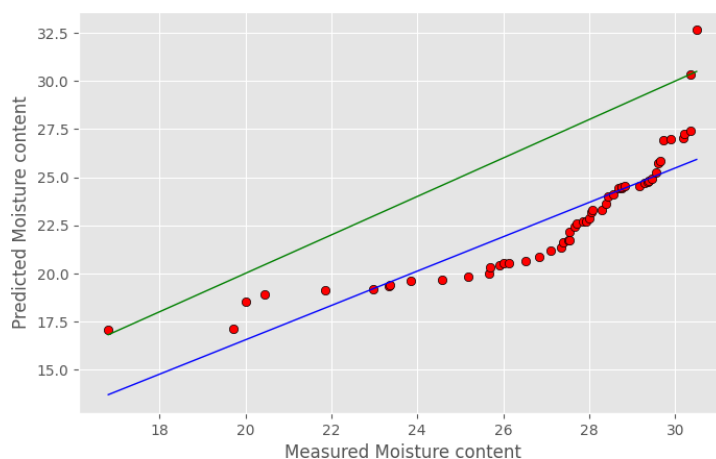


Fig. 4.2.3: Cross-validation of the developed PLS model

This inferred that while the NIR data obtained during the process of soaking are correlated, the moisture content is not much correlated with the spectral values for soaking because the moisture content doesn't change gradually with equal difference with time, there was an abrupt rise at the initial moisture and then very significantly low rise in water uptake. Therefore, the supervised learning technique, principal component regression has been used. The figure on PCA analysis showed the distribution of spectral data from moisture content intervals as PCA only considers the spectral data (and its variance) and not the class labels. Therefore, Principal component regression was used for the purpose of regression on the features reduced by PCA (n=4) (Fig. 4.2.4). The optimal number of components based on accuracy was 4. The prediction accuracy was higher compared to PLS, R^2 of 0.695, and RMSE of 1.109 % for the test dataset (Fig 4.2.5). The standard error (SE) for validation and test set were 3.01 and 3.13 % respectively.

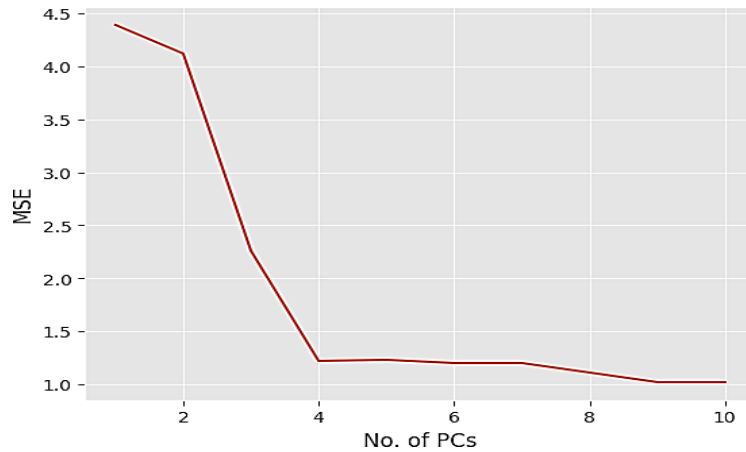


Fig. 4.2.4: Principal Component Analysis

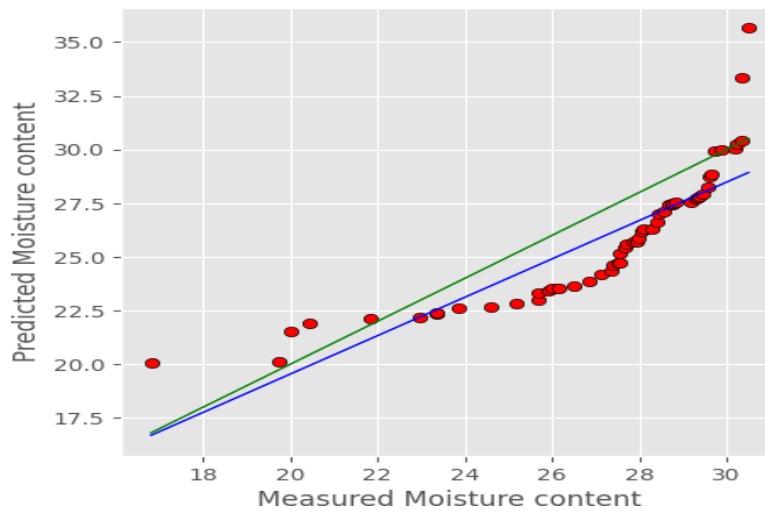


Fig. 4.2.5: Cross-validation plot for PCR

The biases for validation and test set were 0.1 % and -3.8 % respectively. In order to improve the performance of the model, rather than continuous estimation of moisture values during soaking using regression, a discrete approach of classification was chosen. The spectral data were clustered into three different groups of moisture content, (14- 25%), (25-30%), and (> 30 %), and further used for ML-based classification. 0 (14-25%), 1 (25- 30%), and 2(> 30%). A Random Forest (RF) classifier was used to classify the classes based on the spectral information. Therefore, spectral data in the range of 740-1050 nm was used as the features. The features were normalized using standard normal variation (SNV). Previous reports also suggested that ensemble preprocessing gave the best prediction for multivariate spectral data [18]. The dataset of 162 spectra was divided into training (70%), validation (20%), and test set (10%). The model was then trained and validated using k-fold cross-validation (k=5). The best hyperparameters obtained were: 101 number of decision trees in the forest and the maximum depth of a decision tree was 10. The mean classification accuracy was found to be 0.894. The

best model was then used for prediction on the test dataset. The trade-offs between the correctly classified and wrongly classified classes were explained using the confusion matrix (Fig. 4.2.6). The test results represented the probability of each class having been represented in Table 4.4. Since there aren't equal number of classes counts for every class. The class counts for every class as represented in the bar plot (Fig. 4.2.7) showed the class imbalance. Since Class 0, i.e., 14-25% had a lower class count, therefore the Weighted loss functions like weighted cross entropy (CE) (Eq. 3.33) and focal loss (FL) (Eq. 3.34) were measured. The calculated probability for the classes' prediction as in Table 4.4 was incorporated in Eq. 3.33 and Eq. 3.34. The CE value and FL for the minority class (i.e., a class representing 14-25% moisture content) were found to be 0.0079 and 0.0026. Cross entropy is mostly used for classification problems to address the data imbalance problem [19]. Earlier work on NIR spectra of tea leaves for rapid identification of varieties was classified best using LDA and RF classifiers, and the latter performed better [20]. Random forest often performs better because of its ensemble structure that does not lead it to overfitting.

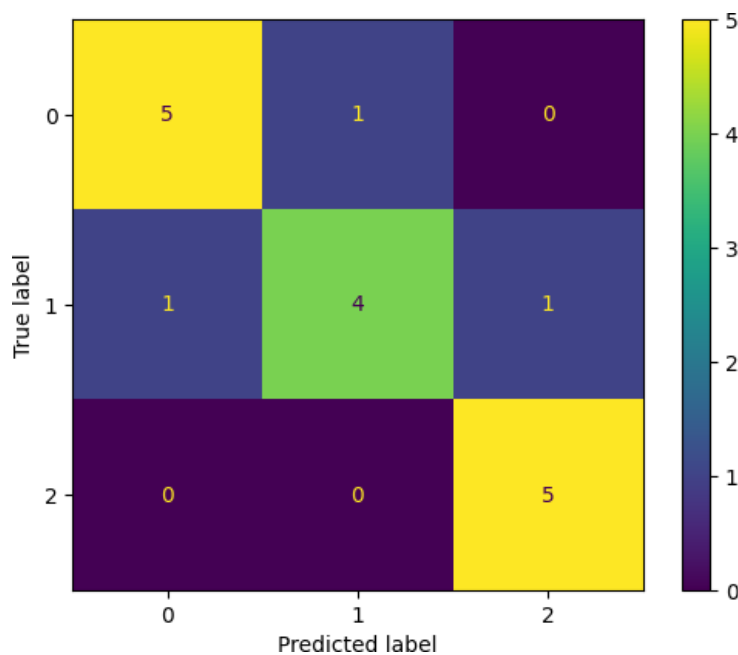


Fig. 4.2.6: Confusion matrix for the prediction classes

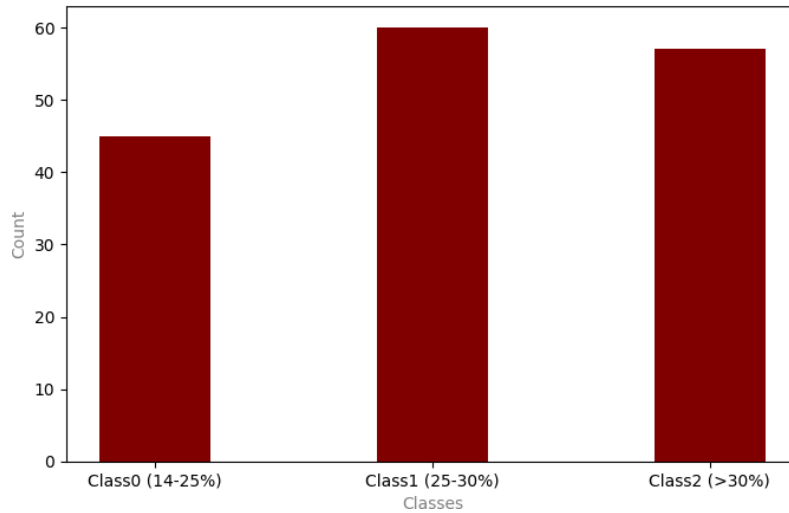


Fig. 4.2.7: Class counts for each class

Table 4.4.1: Probability values of prediction class during soaking

Labels	Class0	Class1	Class2
Class0	0.833	0.2	0
Class1	0.167	0.8	0.167
Class 2	0	0	0.833

4.2.2 Steaming spectral modeling

The raw reflectance values were obtained against the wavelength range for each pressure and time combination in quadruplicate making it a total of 140 scans: over 54 resolution gaps making it 7560 data points. The average reflectance for the range 740-1050 nm was shown for their attributes in Fig. 4.2.8.

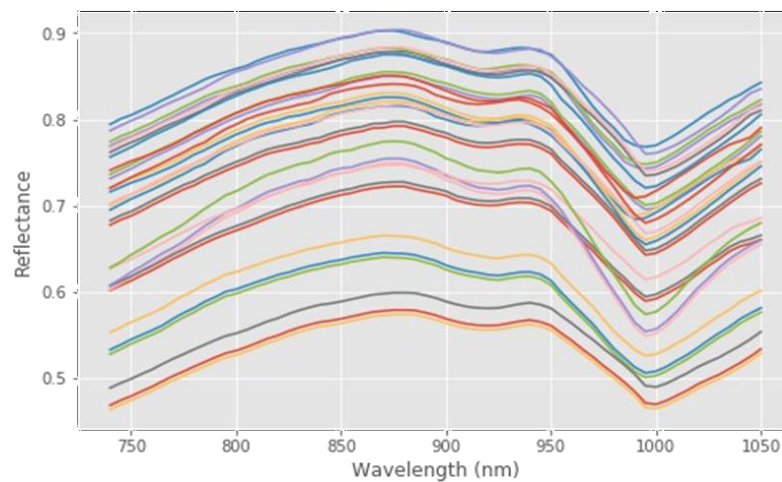


Fig. 4.2.8: Reflectance values against the NIR wavelength range

The raw spectra were pre-processed by using the first derivative function (poly order=1 with odd numbers of windows) to reduce the gaps and to smoothen the spectral data (Fig. 4.2.9). This was followed by logarithmic transformation and the standard normalization.

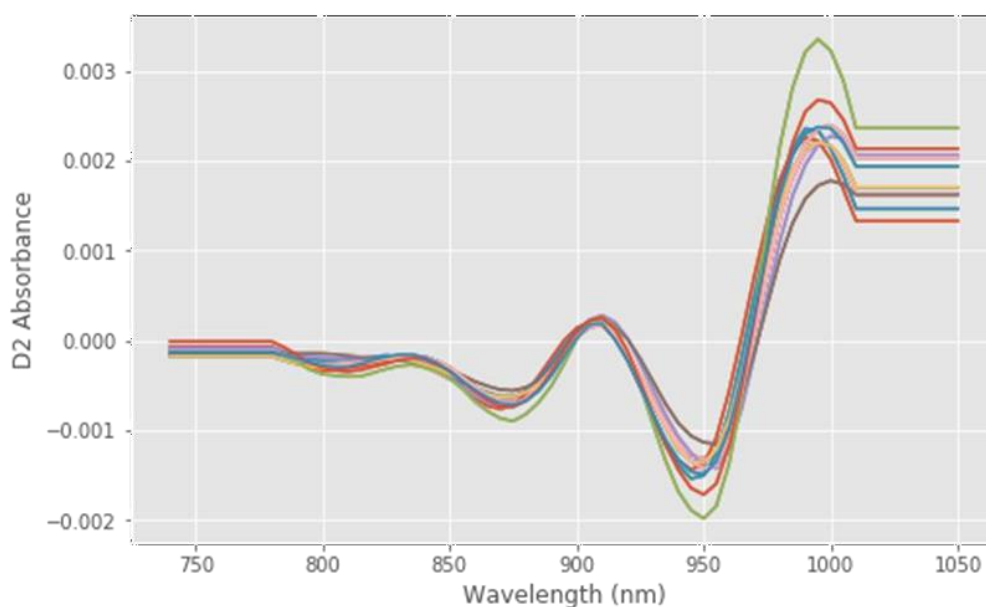


Fig. 4.2.9: Pre-processed spectra of raw reflectance values

Later, the over-scattered range i.e., from 970-1050 nm was kept for the feature set. The feature set was divided into a training set (60%), a validation set (20%), and a test set (20%). The PLS model was fitted on the training set for several components (latent variables). The performance of the trained model was tuned by checking the performance on the validation set. It can be observed that the residual score (RMSE) reaches saturation after 10 latent variables (Fig. 4.2.10). This suggests that 10 latent variables are enough to map the spectrum with their respective DG values. The accuracy for the developed model with 10 LVs in terms of statistical metrics were: R^2 of 0.882 and RMSE of 2.56 % including standard error and biases as 5.16 % and 4.6 % respectively. The precision accuracy of the developed model was assessed on the test set and the predicted values were compared with measured values as shown in Fig. 4.2.11. It can be observed in the cross-validation plot that prediction accuracy at the lower and higher extremes was more accurately predicted than the mid values of DG; thus, showing an R^2 value of 0.843. The estimated highest and lowest relative percent differences (RPD) values were 19.79 and 0.94 %, respectively. The Relative error range (RER) was found to be 12.76 %. There are research reports that give reliability to the use of PLS for the purpose of regression of NIR data for rice characterization. For example, a modified PLS model for optimizing

the NIR reflectance model for measuring the gelatinization characteristics of rice flour [21], amylose determination using NIR [22], and many more [23]. Similarly, portable NIR sensors have given reliable results, for example in the estimation of freshness in quail eggs [24], and tandem eggs [25]; and also, for the classification of ground meat [26] and identifying adulteration in paprika [27]. The present study also found the PLS model equally reliable for modeling NIR spectra with gelatinization degree.

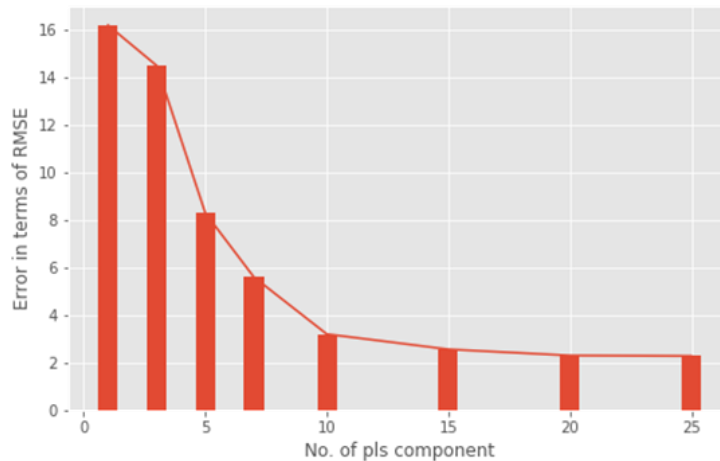


Fig. 4.2.10: Error measurement for different no. of pls components

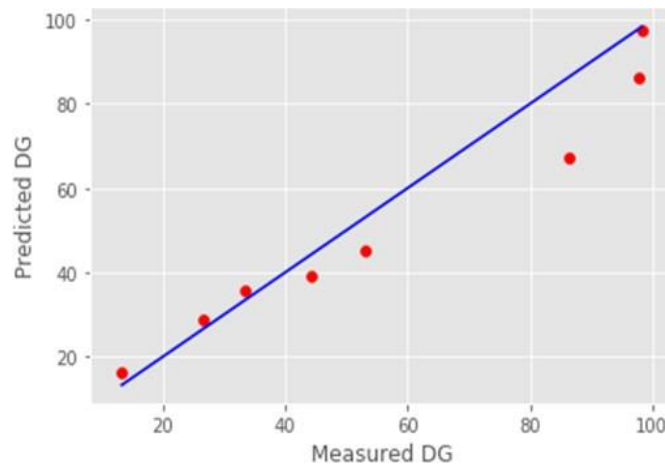


Fig. 4.2.11: Cross-validation plot for degree of gelatinization

4.2.3 Spectral modeling for drying data

The spectra were collected using a laboratory-developed Vis-NIR sensor in the range of 410-940 nm with 18 channels or features (Fig. 4.2.12). The procedure for the development of the Vis-NIR sensor has been mentioned in Appendix-I. The spectral data were collected after every

10 min for 3 h at the 3 temperatures, 40°C, 50°C, and 60°C in a tray drier (Make:UK Model: Armfield – UOP8-A). The number of spectral scans each time was 5, giving a total number of 285 scans. A Neural network using TensorFlow with four hidden layers and the Adam optimizer function was used. The purpose of using ANN over PLS was because there were fewer spectral points or features, and the Neural network is better at finding complex nonlinear mapping with the response labels even if the number of features is less. The dataset was divided into training (70%), validation (20%), and test data (10%). Primarily, the reflectance data set was preprocessed using Z-score normalization. After that, using a TensorFlow sequence of 4 hidden layers with the ReLu activation function, the weights of the neural network were optimized for 100 epochs using an Adam optimizer function. The error loss of the model at every epoch (Fig. 4.2.13) showed that as the epoch increases the mean squared logarithmic error decreases. The validation performance showed a loss of 0.2734. The test score showed R^2 value of 0.624 between predicted and measured values and an RMSE value of 7.034 % (Fig. 4.2.14). The SE and bias are respectively 3.62 and 6.031% respectively. Artificial neural networks have worked well in mapping Vis-NIR data in the food system [28,29].

A classification-based study on the end point screening of rice samples during drying was done using the ML model. The targeted moisture content was kept at less than 13 (% wb) because the best storage moisture percentage is around 12-13%. More spectra of rice samples were taken for samples after drying them to a moisture level of 13% to balance out the number of spectra set into classes to avoid the bias due to inequality of class: Class 0 (< 13%) and class 1 (>13%) moisture content. The same set of data was then trained on the most used classification models (Fig. 4.2.16) without preprocessing as suggested by Li and He [30]. The classification was done based on k-fold validation. The mean accuracy calculated for the validation set was 0.95. The optimal number of hyperparameters chosen was 56 and the depth of the tree was found to be 18. The performance of the test dataset displayed in the confusion matrix (Fig. 4.2.16) showed that the prediction accuracy was 0.933, the recall value of 1, and the precision of 0.87. Therefore, the F-score was found to be 0.928 suggesting a good classification model. A Vis- NIR-based sensor in the range of 400-1000 nm was found to be reliable for estimating moisture in meat earlier [31].

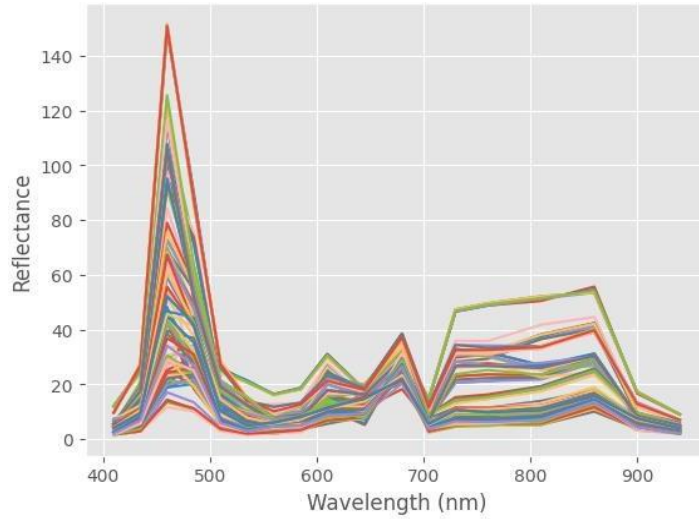


Fig. 4.2.12: Reflectance values against the wavelength

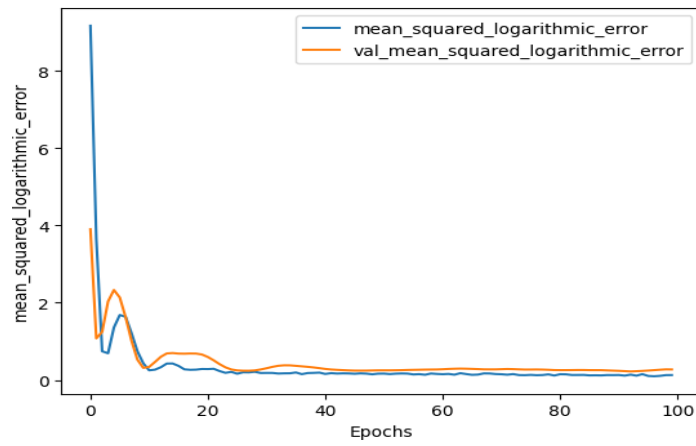


Fig. 4.2.13: Error measurement for ANN

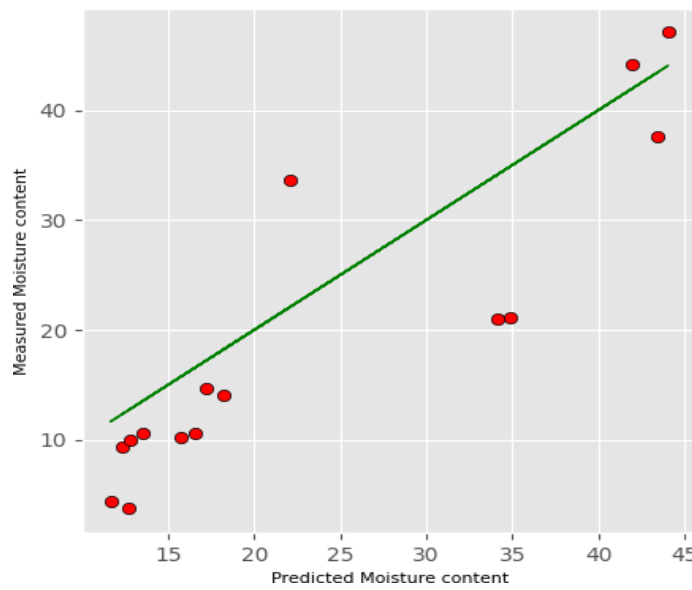


Fig. 4.2.14: Cross validation plot for drying moisture estimation

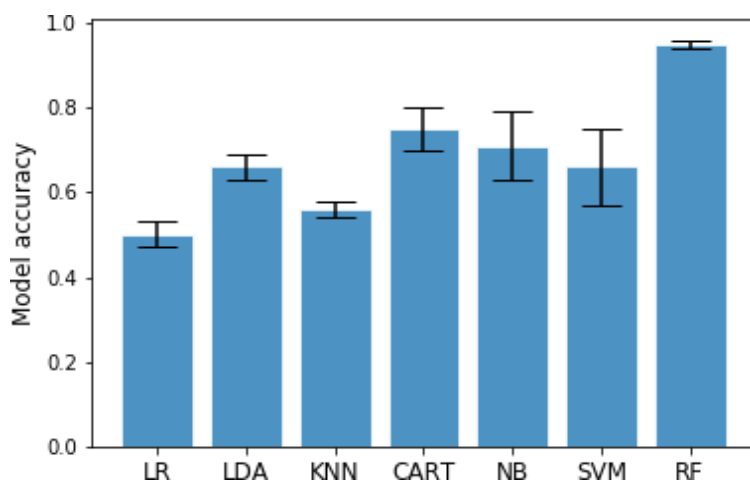


Fig. 4.2.15: Comparison of classification models

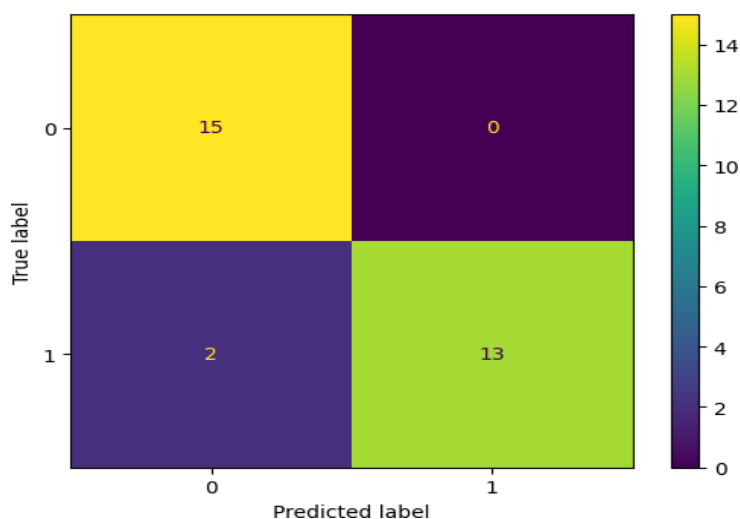


Fig. 4.2.16: Confusion matrix for the test dataset

4.3 Results and discussion for studying the physicochemical changes during ageing

4.3.1 Study on changes in cooking or softening quality of *Komal Chaul*

The change in cooking or softening quality of *Komal Chaul* with respect to storage time in a year showed that the cooking time had increased linearly (Fig. 4.3.1). The starch granules absorbed water and caused the starch to gelatinize, and cooking time was considered as the time when all starch (> 90%) had gelatinized. The low amylose composition and the soft nature of this rice causes it to cook faster. There was an increase in cooking time with respect to ageing. The slight inconsistency of the linear curve might be due to the difference in the extent of processing while parboiling the *Chokuwa* rice, though similar conditions were maintained for every batch, the slight difference may be shrinkage of the surface of certain grain during

drying. The change of physical parameters like true density from $1.26 \times 10^3 \text{ kg/m}^3$ to $1.37 \times 10^3 \text{ kg/m}^3$ and bulk density from $0.857 \times 10^3 \text{ kg/m}^3$ to $0.897 \times 10^3 \text{ kg/m}^3$. The reason for this could be due to the increased air spaces or porosity as the rice dries out due to syneresis or retrogradation of *Komal Chaul* that might have caused the grain to become compact resulting in less absorption of water and as a result higher cooking time for the aged sample (Appendix-III).

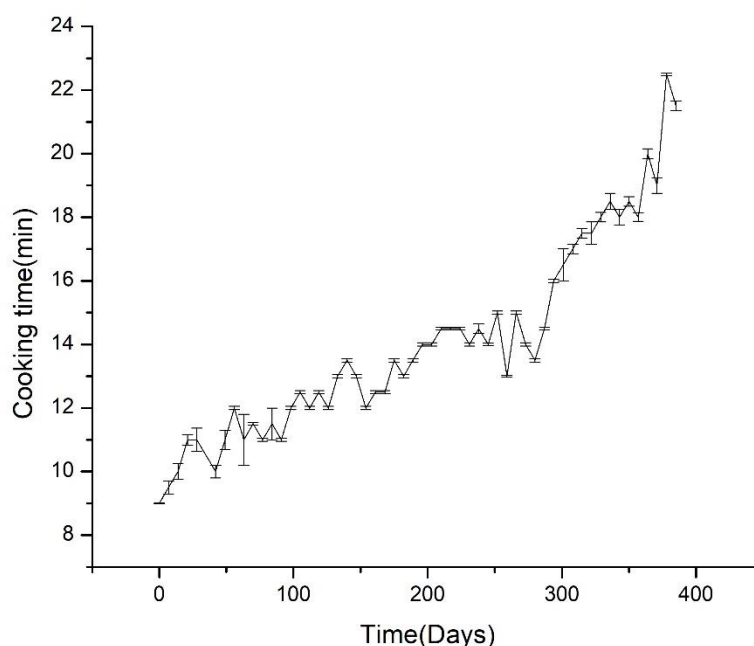


Fig. 4.3.1: Cooking behavior with respect to storage time

4.3.2 Changes in composition

The rice samples were marked as UANP: Unaged Non-Parboiled, UAP: Unaged Parboiled, ANP: Aged Non-Parboiled, AP: Aged Parboiled. The respective amylose content obtained for the samples UANP, UAP, ANP, and AP were 9.70 ± 0.29 , 9.10 ± 0.25 , 12.72 ± 0.07 , 12.39 ± 0.57 %. The amylose content or the apparent amylose content showed no significant change during the process of ageing. There was a significant difference in the percentage of amylose in parboiled and non-parboiled samples (Fig. 4.3.2). This was because of the formation of short amylopectin fine structures during the process of parboiling.

The protein content for the samples was estimated for both aged and unaged rice (Fig. 4.3.3). The aged samples (ANP) and unaged samples (UANP) of *Chokuwa* showed significant difference. The protein of *Komal Chaul* (UAP) and (AP) showed differences too. Oryzenin is a

chief peptide in rice. The decrease in the concentration of low molecular weight protein in the aged rice grains might have resulted in a change in the protein concentration [31]. A different study also reports an increase in the peptides of higher molecular weight like globulins during storage [32].

The fat content for samples UANP, ANP, UAP, and AP were 2.27 ± 0.02 , 2.03 ± 0.02 , 0.56 ± 0.03 , 0.55 ± 0.05 g/100g respectively (Fig. 4.3.4). The fat content in the parboiled sample was much lesser because of inhibition of lipase activity during processing. The free fatty acid value increased from 0.118g /g (UANP) to 0.253 g/g (ANP) for *Chokuwa* rice. While for *Komal Chaul*, free fatty acid increased from 0.0595 g/g rice to 0.0613 g/g. As an earlier proposed theory by Moritaka and Yasumatsu as stated by Zhou et al. [33], the consequences of hydrolysis of lipids to free fatty acids result in forming a complex between starch and interaction with starch and protein.

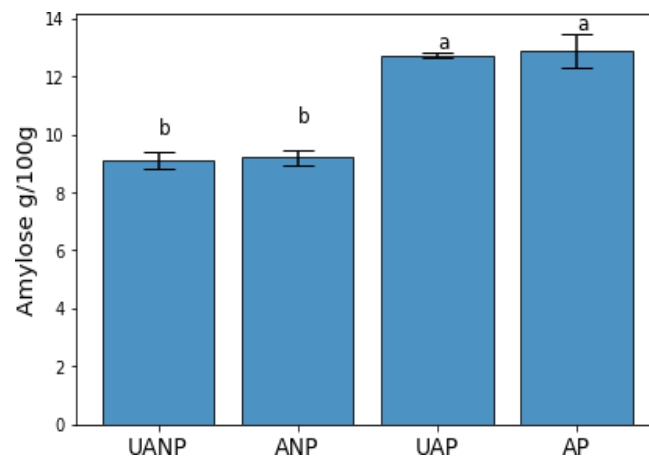


Fig. 4.3.2: Amylose content in aged and unaged rice

(Different alphabets represent significant differences based on Duncan Test, $p < 0.05$)

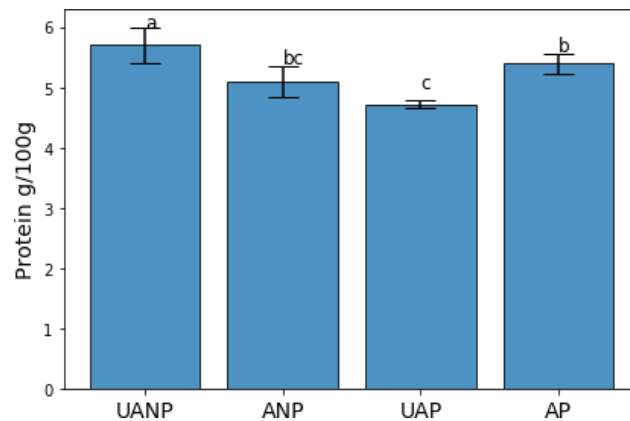


Fig. 4.3.3: Protein content in aged and unaged rice

(Different alphabets represent significant differences based on Duncan Test, $p < 0.05$)

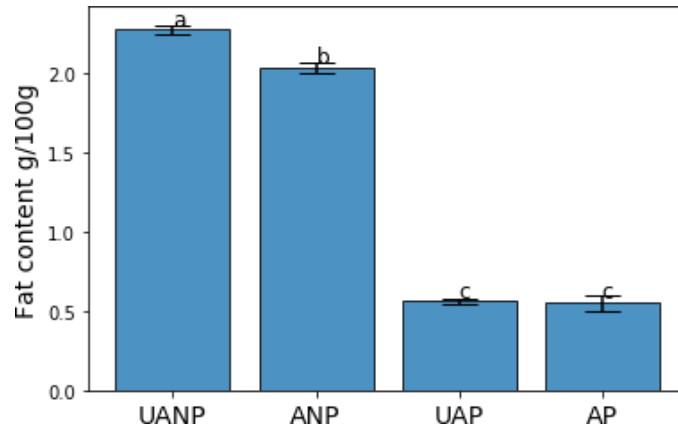


Fig. 4.3.4: Fat content in aged and unaged rice

(Different alphabets represent significant differences based on Duncan Test, $p < 0.05$)

4.3.3 Changes in textural parameters

Textural parameters were measured for aged and unaged parboiled rice after rehydration and showed variation in hardness and stickiness. The overall average change in the value is quite significant (Fig. 4.3.5). For the aged samples, hardness values were higher while stickiness values were much lower compared to that of cooked unaged parboiled rice. The increase in hardness could be due to grains becoming denser and more compact as rice loses moisture and the gaps between the starch granules might have decreased over time [34].

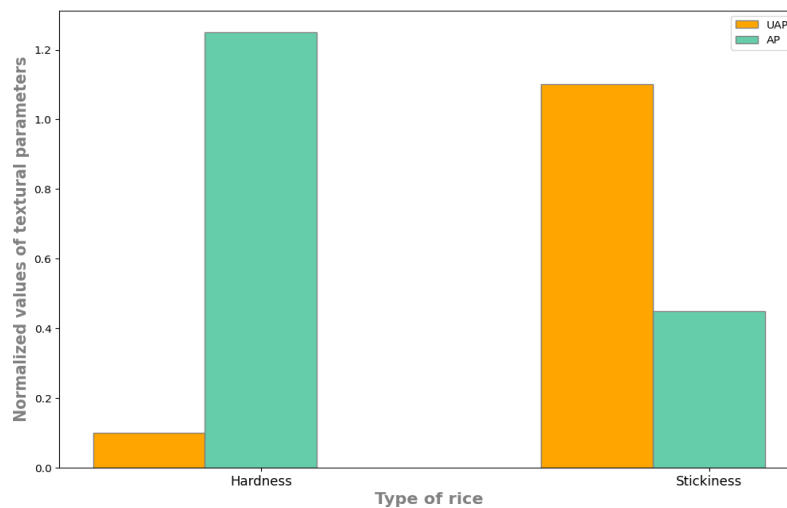


Fig. 4.3.5: Textural parameters of aged and unaged rice

4.3.4 Changes in pasting properties

There were abrupt changes noticed in the pasting properties of the aged and unaged samples (Fig. 4.3.6). For non-parboiled samples peak viscosities and setback viscosities have changed much over a year. However, the pasting temperature of parboiled samples was lower indicating

their no cooking properties (62-64 °C). Parboiled samples (*Komal Chaul*) shifted to lower values of peak viscosity, holding viscosity as well as final viscosity. The viscosities such as the final and setback viscosity of aged rice usually increase [34]. For non-parboiled rice i.e., *Chokuwa* rice, a similar trend was observed. The final viscosity, setback viscosity, and peak viscosity were significantly higher for ANP than UANP. The properties of aged *Komal Chaul* showed lower setback however the setback and holding viscosities were higher than the unaged parboiled sample (UAP).

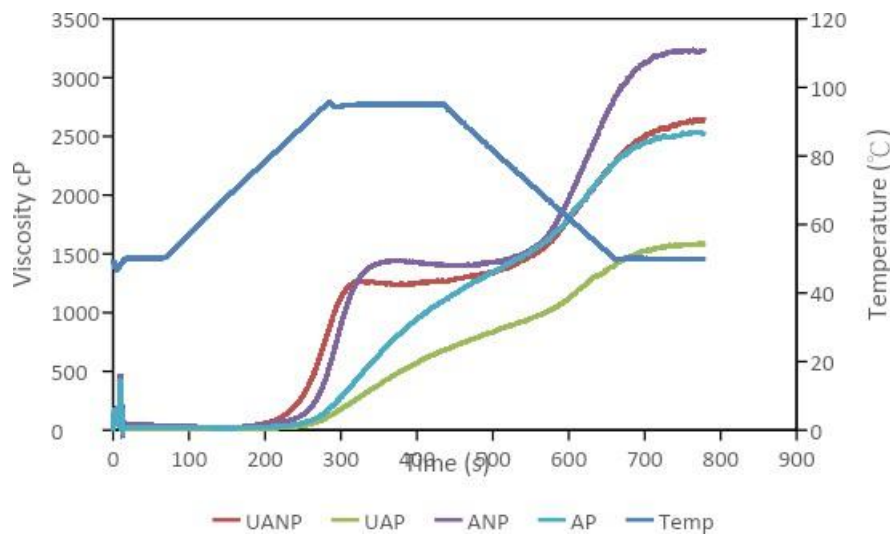


Fig. 4.3.6: RVA plot of aged and unaged rice

4.3.5 Changes in molecular bonds

The FTIR plots of *Chokuwa* rice showed strong OH, C-H, and C=O peaks at 3200-3400 cm^{-1} , 2850-3100 cm^{-1} , and 1650-1800 cm^{-1} respectively (Fig. 4.3.7). The ANP spectrum obtained is quite similar to the unpolished raw sample of *Chokuwa* used for fortification, this may be due to the presence of cellulosic materials or bran layer in the aged rice [35, 36]. The freshly harvested *Chokuwa* rice had the same peaks while the bond intensity at 3200-3400 cm^{-1} is lesser than the aged sample which could have changed due to the interaction between starch protein or starch lipid interaction. For *Komal Chaul* there were no distinguishable peaks observed for OH and CH stretching. Similar peaks were found for parboiled *Komal Chaul* [36].

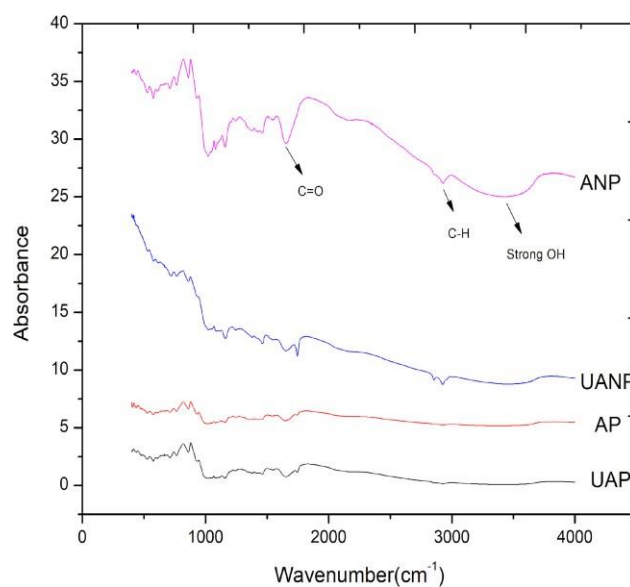


Fig. 4.3.7: FTIR plot of aged and unaged rice

4.3.6 Changes in crystalline pattern

The crystallographic change in the aged (A) and unaged (UA) sample of *Chokuwa rice* (NP) and *Komal Chaul* (P) were observed using X-ray diffraction plots (Fig. 4.3.8).

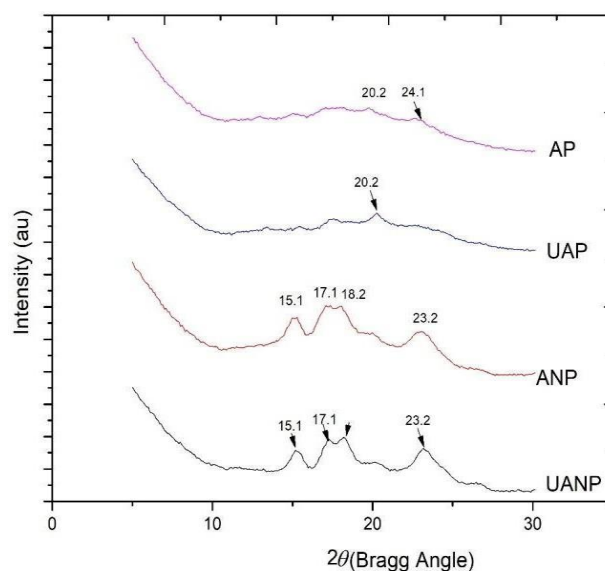


Fig. 4.3.8: XRD plot of aged and unaged rice

The *Chokuwa rice* samples both aged (ANP) and unaged (UANP) showed peaks at 15.1°, 17.1°, 18.2° and 23.2° which suggests A-type diffractions as mentioned by Dutta and Mahanta for

Chokuwa rice [37]. The peak widths have increased for ANP and as a result, the crystallinity percentage was found to increase from 36.9% to 47.5%. *Komal Chaul* showed an almost amorphous diffraction plot with a feeble V type at 20.2° Bragg angle for the unaged sample (UAP) contributing to a crystallinity percentage of 13.23%. An additional B (24.1°) type peak was observed in the aged sample of *Komal Chaul* (AP) resulting in an increase in crystallinity percentage from 13.23 to 19.24% which could be inferred as the cause of higher cooking time in the aged sample. Similar peaks of *Komal Chaul* were observed for pressure-parboiled samples for longer duration [38].

4.3.7 Temperature-based storage study of parboiled *Komal Chaul*

4.3.7.1 Effect on rehydration property

Water absorption or rehydration ratio is an index for the efficiency of cooking quality. Higher absorption suggests better cooking quality [39]. There were significant changes that had been observed in the rehydration ratio of soaked parboiled *Komal Chaul*, especially for the sample stored at 37 °C (Fig. 4.3.9).

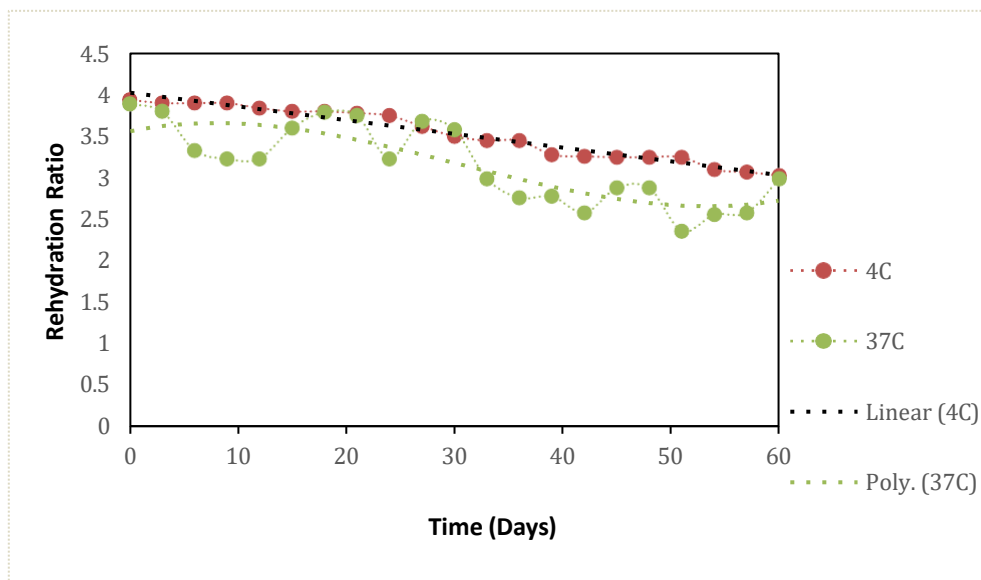


Fig. 4.3.9: Water uptake behavior of sample stored at 4 °C and 37 °C

The lowest rehydration ratio was observed at 3.45. The sample stored at 4 °C showed negligible variation in the water uptake behavior till and the 24th day of storage (Fig. 4.3.9), and on the later days, there was minimal change in the ratio. Statistical comparison using two paired t-tests showed that the most significant change with P-value ($T \leq t$) less than 0.01, the mean difference of rehydration ratio was observed to be more in the rice sample stored at 37 °C. Earlier

studies, on the other hand, it reported that rice varieties stored at 37 °C showed greater water uptake, reduced pH, and solid loss, however, those rice samples were high and intermediate amylose rice with amylose content in the range of 29-18% [40]. Because no study on the rehydration property of no cooking low amylose variety *Komal Chaul* concerning ageing had been done, therefore, a conclusive remark requires further analysis. The change in the water uptake behavior could be possibly due to chemical or physical changes in the structure during storage. The low temperatures might have not facilitated much change, but at higher temperatures the changes were distinguishable.

4.3.7.2 Effect on pasting property

Forty viscograms were obtained from *Komal* rice stored at 37 °C and 4 °C for a period of 2 months at an interval of 3 days (Fig. 4.3.10). For the purpose of representing and finding the differences, the RVA data were analyzed using PCA as shown in Fig. 4.3.11. The two-component PCA analysis (Fig. 4.3.11) showed that samples stored at 37 °C for 33, 42, 45, 54, 57 days (i.e., the labels 37_33, 37_42, 37_45, 33_54, and 33_57) showed less covariance from rest. Samples stored at 37 °C, those profiles when compared with the 0th-day sample it can be observed that the samples for the latter days' viscosities shifted to lower values. There was an increase in peak viscosity, breakdown viscosity, setback as well final viscosity for a few days, later for days, and then a sudden decrease from the 27th day onwards. The pasting temperature plot gave a clear observation of how the pasting temperature varies with storage time. For *Komal Chaul* stored at 4 °C, there was no significant difference observed in the pasting temperature throughout the storage period of 60 days having an average temperature of 61.2 °C (Fig. 4.3.10). However, in the pasting temperature profile for P_37 (*Komal Chaul* stored at 37 °C), there was more variation in the temperature plot ranging from 52-75 °C approximately.

Thus, if the storage temperature was high, there was more variation in viscogram profiles (Appendix-IV). Huang and Lai [41] used multivariate factor analysis, Principal Component Analysis (PCA) for the representation of RVA profiles. Rice stored at low temperature exhibited a lesser degradation as cited in earlier literature; including intensive degradation at high storage temperature and delayed storage period were reported for waxy rice stored at 4 and 17 °C [40, 41].

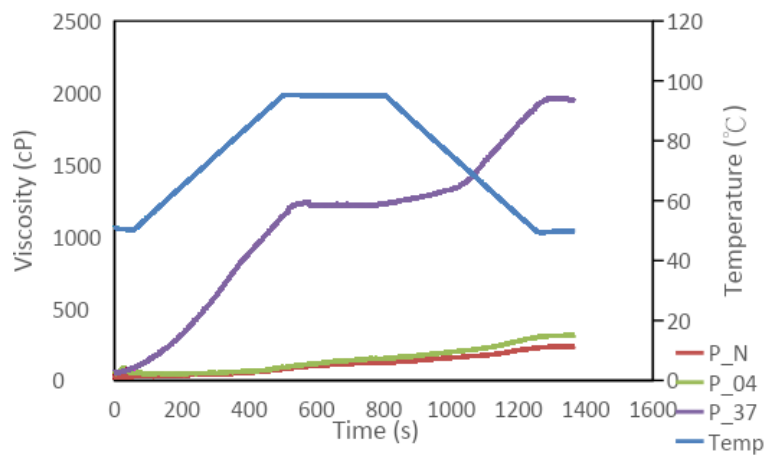


Fig. 4.3.10: Viscograms at different at different temperature and time

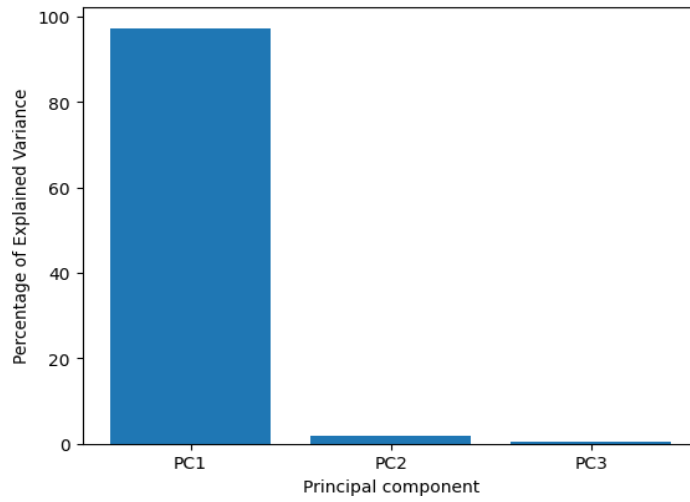


Fig. 4.3.11: PCA scree plot of viscogram data

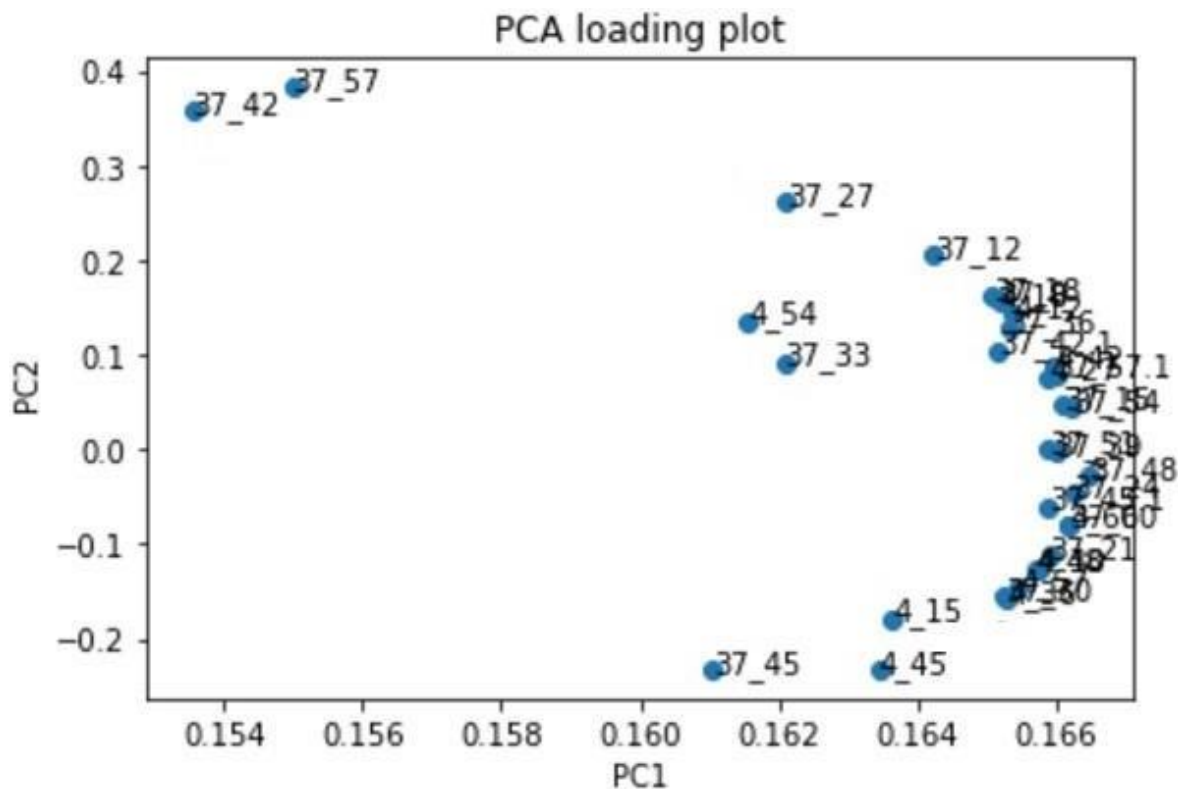


Fig. 4.3.12: PCA biplot of viscograms at different temperature and time.

(The interpretation of these two number codes is: first numbers of 4 and 37 are the storage temperatures in degree Celsius, second number is the number of days of storage of samples).

4.3.7.3 Effect on thermal property

Thermal characterization was done using the shift in peak temperature as well as the change in peak width of the endothermic curves were quite distinct. *Komal Chaul* stored at 37 °C (P-37) had a much broader and steeper endothermic peak (Fig. 4.3.13). The onset temperature and peak temperature had also changed considerably. The broader, as well as sharper peak at 37 °C storage, indicates that the rice took more time and thermal energy to gelatinize; this result was substantially attributed to the partial softening of starch at a warm temperature and then reordering of the starch granules of parboiled *Komal Chaul* during the process of ageing. *Komal Chaul* stored at 4 °C (P-04) showed a shift in onset temperature i.e. 46.3, which is more than the sample P-37. P-04 may not have caused any further degradation or reorientation of the starch components, and similar endothermic peaks were obtained for mildly parboiled *Komal Chaul*. The polymers in P-04 may have similar temperature sensitivity and that is the reason why it has merged with the main peak, which suggests the retrograded fraction [42]. This inference of DSC thermograms ratifies with our structural morphology and crystallinity values.

No peaks were obtained after 100 °C for amylose lipid complexes. Consequently, previous literature suggests that high percentage of amylopectin facilitated the starch granule structure to be stable and which caused a rise in the gelatinization temperature and enthalpy determined by differential scanning calorimetry.

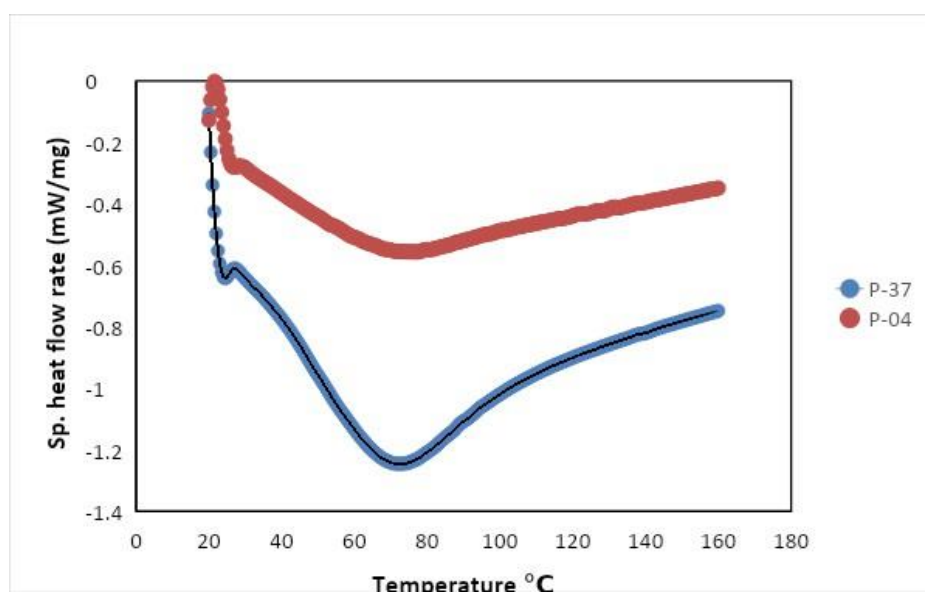
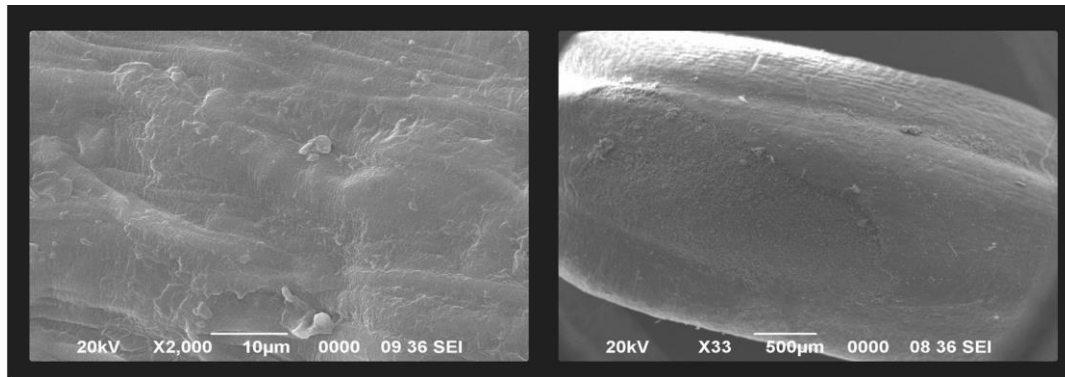


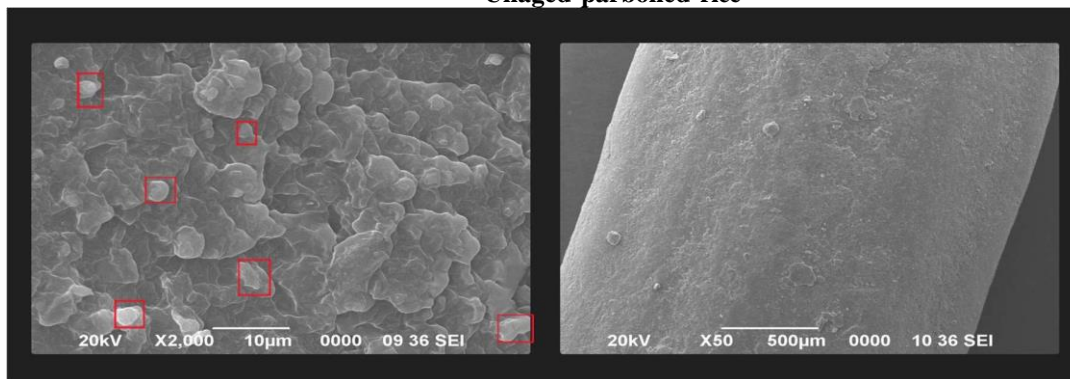
Fig. 4.3.13: Differential Scanning Calorimeter plots at storage temperature of 37 °C and 4 °C

4.3.7.4 Effect on surface morphology

SEM was used to visualize the internal structure of unaged and aged parboiled rice grains and to examine the effect of the microstructure of individual parboiled rice grains due to temperature storage. The unaged parboiled rice from the freshly harvested sample shows a very smooth surface, while in contrast, the aged rice's SEM image for both 37 °C and 4 °C showed globular structure in the images taken at the highest focal length (Fig. 4.3.14 and Fig. 4.3.15). It can be considered that the hydrogen bonds between the amylopectin side chains played an important factor in the growth of the structures, and the amount of water present both have a significant impact on the internal structure of the granules [42,43]. There was no round structure in the unaged SEM image, samples at 4 °C showed formations of a few tiny round structures. The average diameter of a few larger-sized globular starch structures (marked red in Fig. 4.3.15) was found to be 9.6 μm, quite high as compared to the average diameter of rice stored at 4 °C. The appearance of a white line might have been attributed to the white swollen globular structure in the magnified image which could be possibly due to partial gelatinization of parboiled *Komal*

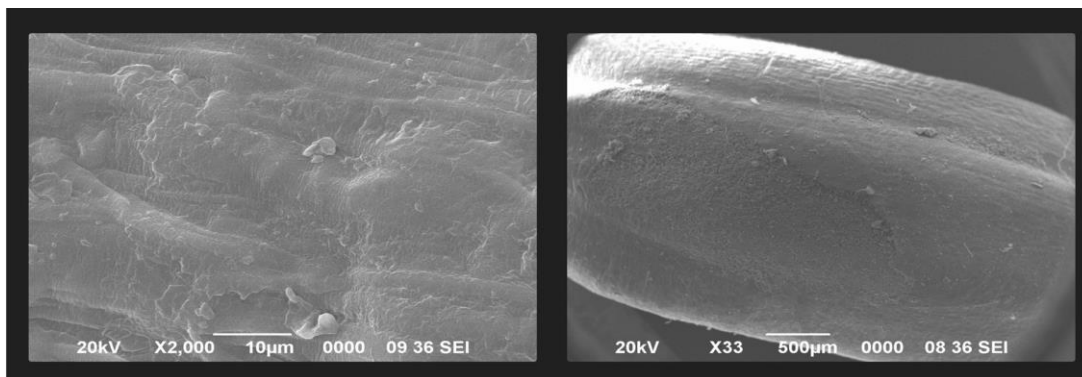


Unaged parboiled rice

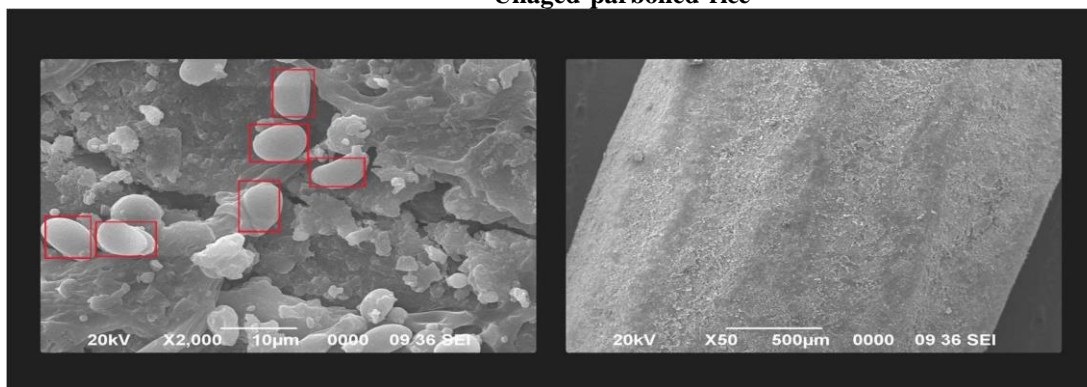


4 °C storage

Fig. 4.3.14: Comparison SEM images of *Komal Chaul* after lower temperature image



Unaged parboiled rice



37 °C

Fig. 4.3.15: Comparison SEM images of *Komal Chaul* after higher temperature storage

Chaul at a warm temperature, and then settling of the partially gelatinized granular particles when brought to room temperature. The other reason could be the interaction of starch protein molecules which was one of the reasons cited earlier on the study of waxy rice of protein.

4.4 Results and discussion for prediction of ageing time

4.4.1 Estimation ageing time of *Komal Chaul* by ML model

Firstly, NIR-based spectral library (wavelength range: 740-1050 nm) of aged and unaged rice samples with their respective ages from the time of harvesting was created (Fig. 4.4.1). Those raw data were preprocessed using the Savitzky Golay function, that takes the derivative of the data (Fig. 4.4.2). Then, a Partial Least Square (PLS), an ML-based regression model was trained on the pre-processed data, validated, and tested (k-fold selection) for estimating the time of ageing of rice using the spectral library. Later, the predictive model with best-fitted parameters. The least MSE values were obtained for 10 latent variables, so a PLS model with 10 components was trained (Fig. 4.4.3). The PLS model for ageing time estimation showed R^2 of 0.897 and RMSE of 19.409 days for the validation set. The test dataset showed a lesser accuracy with an R^2 value of 0.7013 and RMSE to be 32.477 days (Fig. 4.4.4). However, the relative percent deviation was less, that is 3.08.

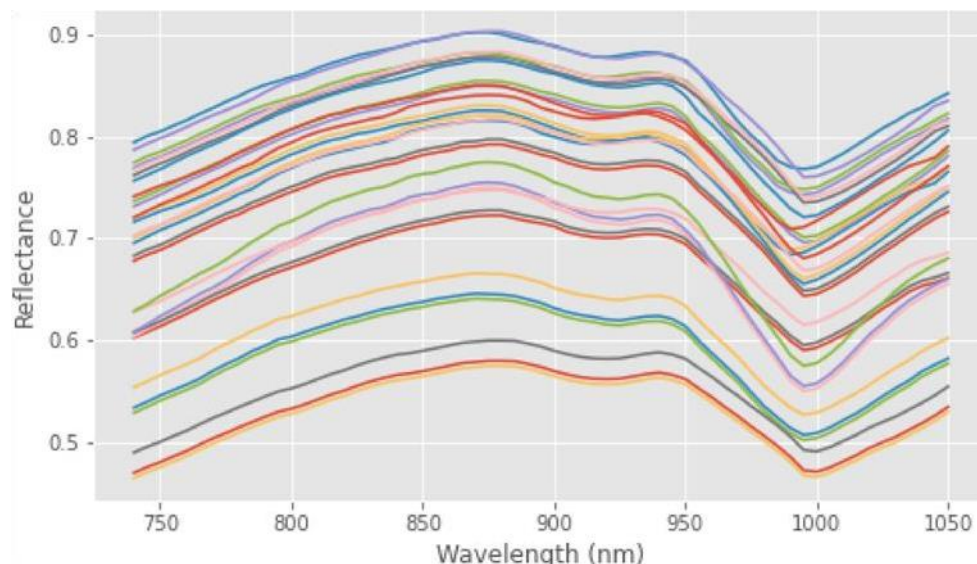


Fig. 4.4.1: Raw reflectance vs wavelength

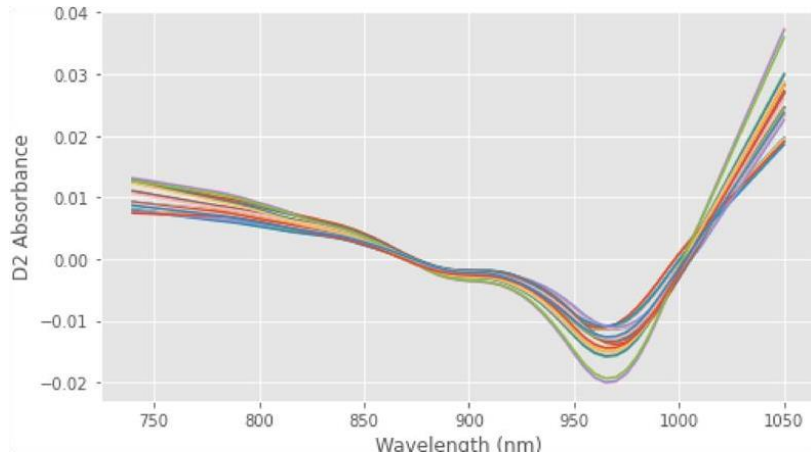


Fig. 4.4.2: Pre- processed spectra of raw reflectance

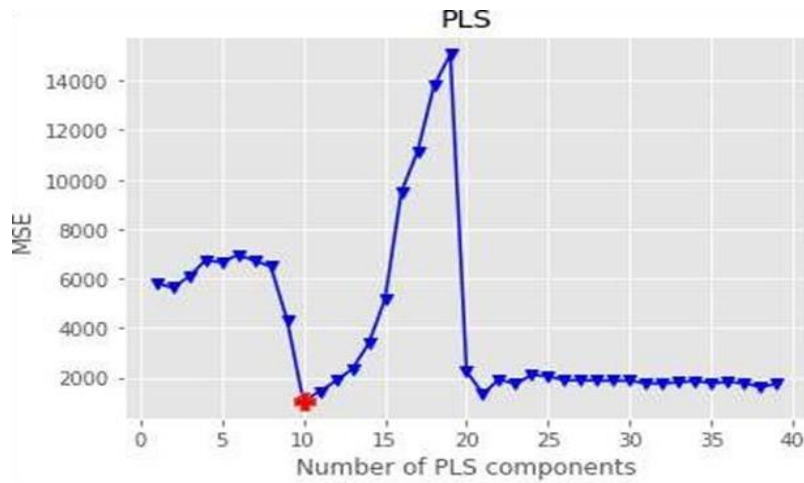


Fig. 4.4.3: Hyper-tuning the Latent variables

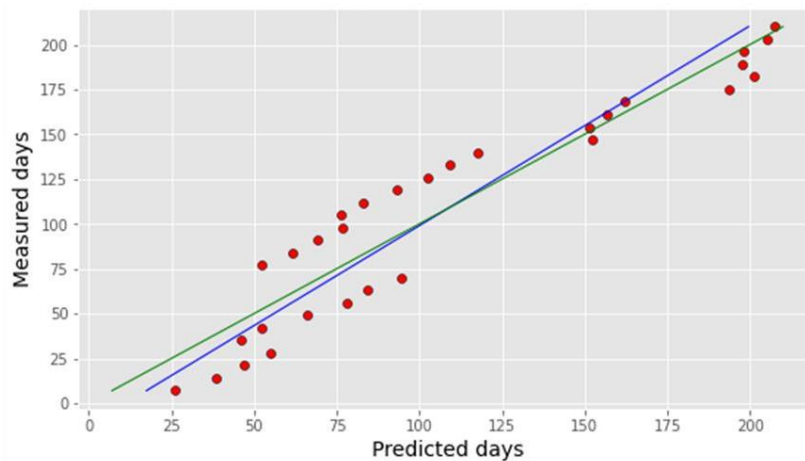


Fig. 4.4.4: Cross validation plot of test data for ageing time

To improve the performance of the model a variable selection approach was made. Earlier studies had suggested that variable selection for high-dimensional data was helpful in improving the performance of those with higher dimensions of data. There are various ways to approach selecting variables for PLS regression: filter method, wrapper method, and embedded method. For this study, a filter method was applied for wavelength selection [44]. The filtration was based on the regression coefficients of the PLS model. The association between PLS coefficients for every wavelength gave us an idea of largely contributing features (Fig. 4.4.5).

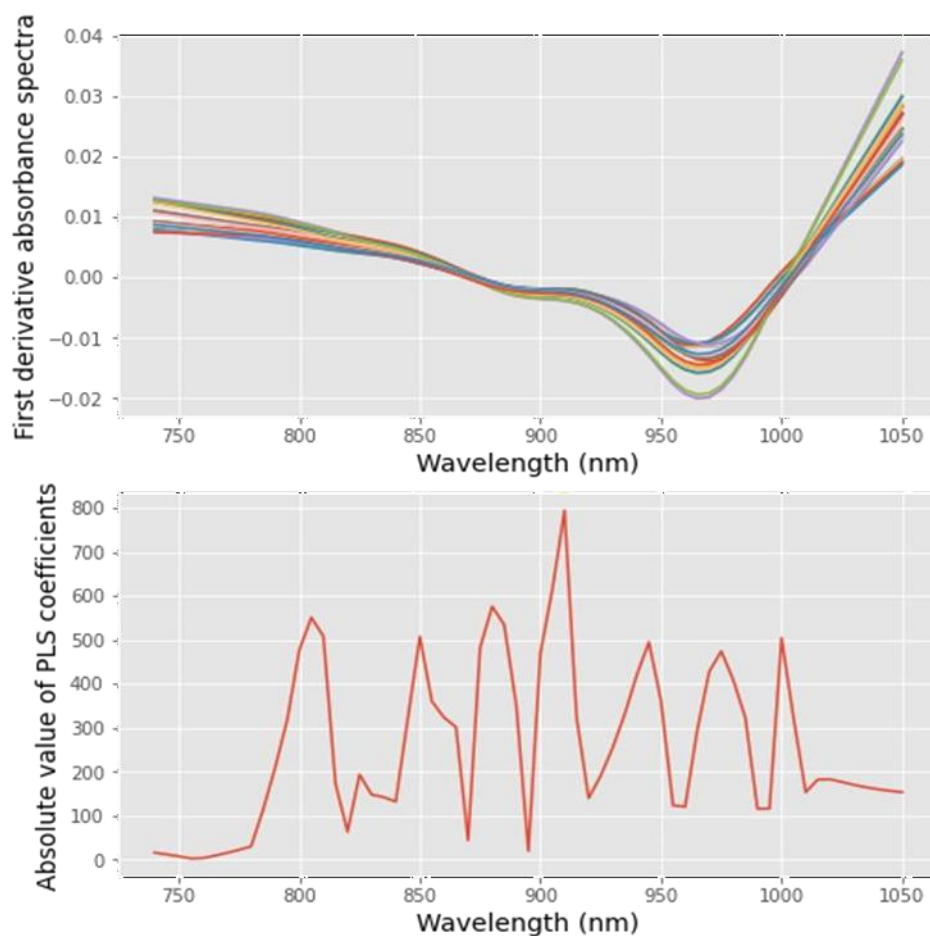


Fig. 4.4.5: Fitting parameters for optimizing features

The number of PLS components after optimization was chosen to be 40. The application of a recursive method whereby eliminating one wavelength at a time and calculating the MSE value. The optimal number of discarded wavelength features was 37 based on an optimal MSE value of 11.29. The chosen wavelength values were 790, 795, 800, 805, 810, 830, 835, 840, 845, 850, 855, 860, 875, 880, 885, 900, 905, 910, 930, 935, 940, 945, 970, 975, 980 and 1000 (Fig. 4.4.6).

The cross-validation plot between the measured and predicted showed better regression with an R^2 value of 0.89 (Fig. 4.4.7). The regression for the test dataset was better than the previous PLS model suggesting better predictability. The RMSE values of validation and test set were found to be 2.01 and 3.24 days which is considerably low. The SE for validation and test set were found to be 1.75 and 14.24 days and biases were 3.9 and 9.6 days respectively.

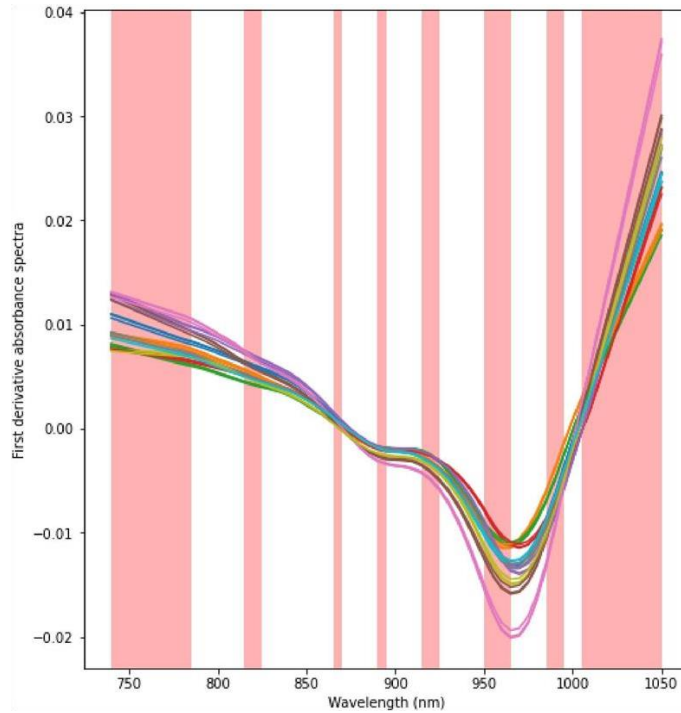


Fig. 4.4.6: Optimal features selection

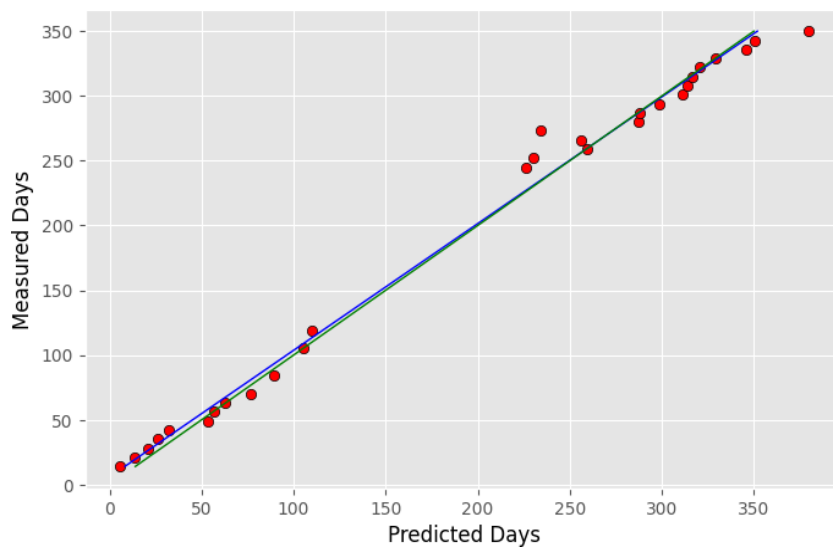


Fig. 4.4.7: Cross validation plot after optimization

4.4.2 Estimation of cooking quality of *Komal Chaul* by ML model

The same spectral data (Fig. 4.4.8) were later mapped with the cooking time or softening time of *Komal Chaul* have changed linearly with time. Those raw data were preprocessed using the Savitzky Golay function, which takes the derivative of the data. PLS regression was used to model cooking with spectral data taken during the process of ageing. The training set was trained over 40 numbers of Latent variables. The least MSE values were obtained for 10 latent variables, so the PLS model with 10 components was validated. The validation score was found to be R^2 : 0.79 and RMSE of 1.09 min. The test dataset showed accuracy with an R^2 value of 0.77 and an RMSE to be 2.77 min (Fig. 4.4.9). The RMSE was higher because the endpoint predicted deviated more for both the higher values and lower values.

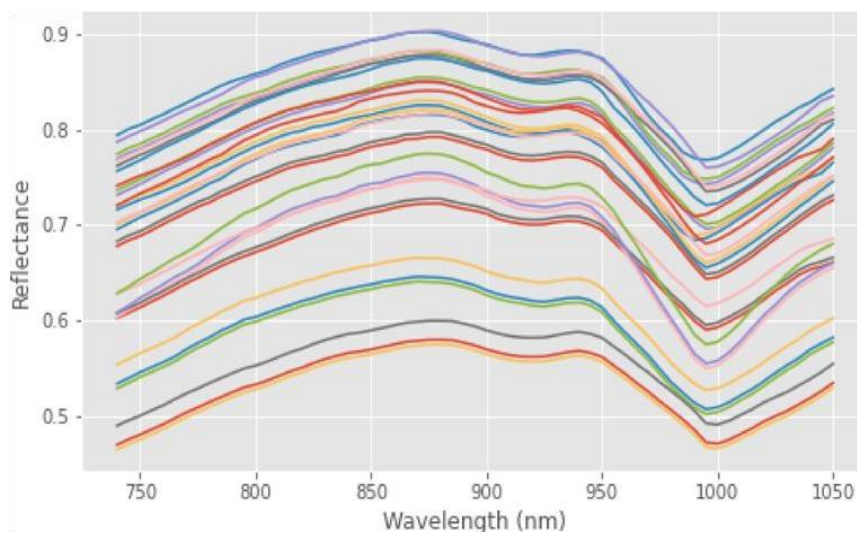


Fig. 4.4.8: Raw reflectance vs wavelength

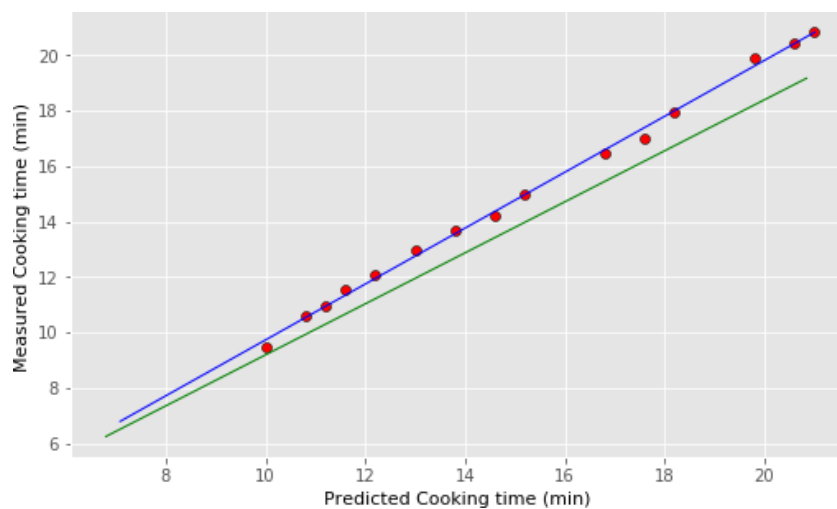


Fig. 4.4.9: Cross validation plot for prediction of test data for cooking time

Therefore, ML-based PLS regression method proved quite a definitive technique in calibrating the ageing parameters of *Komal Chaul* as supported by previous work that worked on aged rice determination using NIRS [45]. This is because, in the case of limited data, PLS is more stable than other models as it can establish non-linear relationships between input and output variables by exploiting underlying latent factors.

4.4.3 Prediction of the age of *Komal Chaul* by classification ML model

The ageing data were collected over a period of a year at an interval of 7 days. The spectral archive of ageing data for *Komal Chaul* was divided into 3 groups: 1 month, 3 months interval, 6 months interval. The classification models were then trained on the training dataset and tested over the test dataset using the k-fold validation technique. The validation accuracy was found to be highest for the SVM model (Fig. 4.4.10) with a value of 0.77. This could be due to ability of SVM models to form collinearity among multilinear data [46]. The test accuracy was found to be 0.79 and the F1-score was found to be 0.69. The test dataset predictions were represented using a confusion matrix which showed that the later month predictions are coinciding (4th to 7th group) with that of the 6–9 months predictions. (Fig. 4.4.11). Therefore, our next classification was based on three months classification, and the random forest classifier showed the best in classifying the data where each group consisted of spectral data from 3 months. The test accuracy was found to be 0.89 and the F1 score was found to be 0.81. This showed spectral data from 6-9 months (2nd group) coincides mostly, also the 0th and 1st group suggesting lower accuracy of prediction for the first 3 months. Lastly, a classification was based on six-month spectral data. The prediction accuracy was highest for RF classifier with an accuracy of 0.92 for the test data and the F1 score was highest. Therefore, the classification comparison suggests that *Komal Chaul* that are spectra obtained from apparently age groups of 6 months, were showing better predictability. This could also be due to the reason of a two class classification problem.

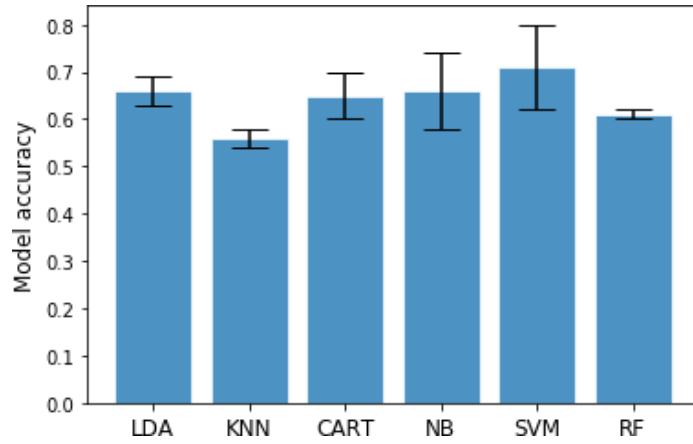


Fig. 4.4.10: Comparison of classification of ageing time on 1-month basis

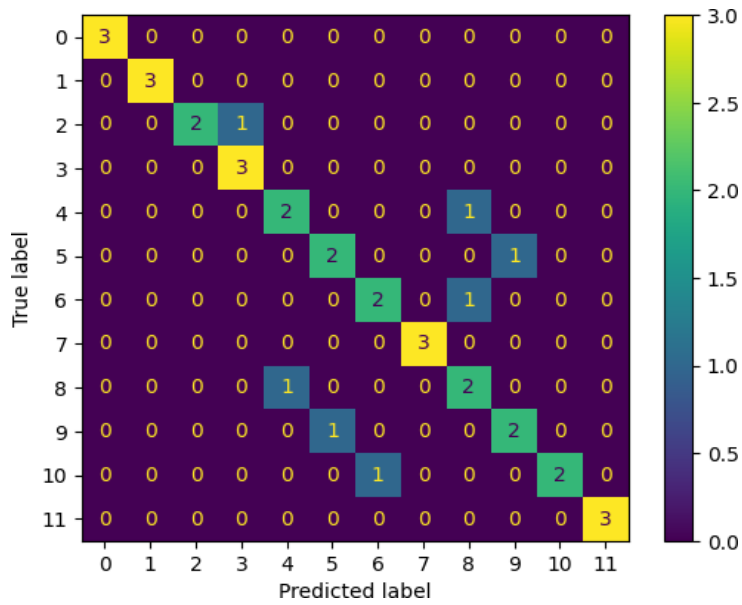


Fig. 4.4.11: Confusion matrix of classification of ageing time on 1-month basis

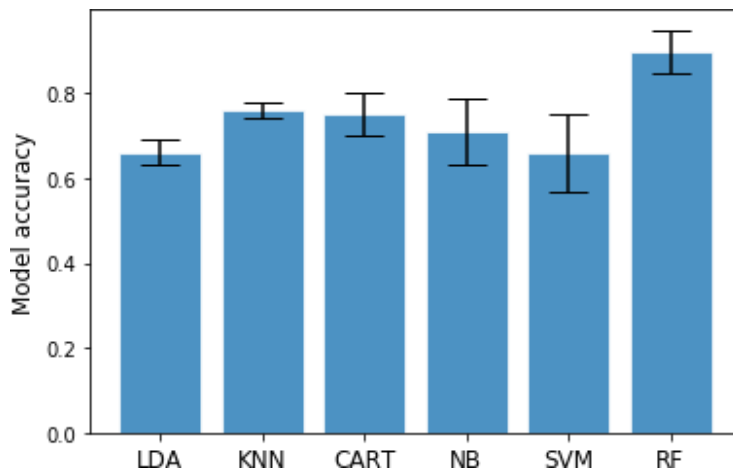


Fig. 4.4.12: Comparison of classification of ageing time on 3-month basis

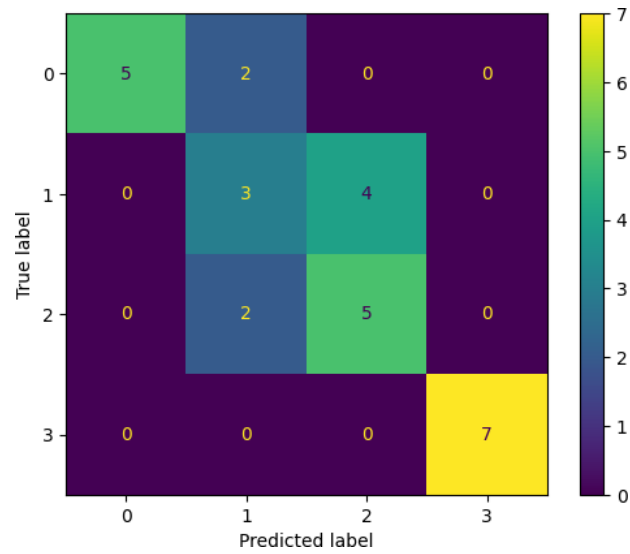


Fig. 4.4.13: Confusion matrix of classification of ageing time on 3-month basis

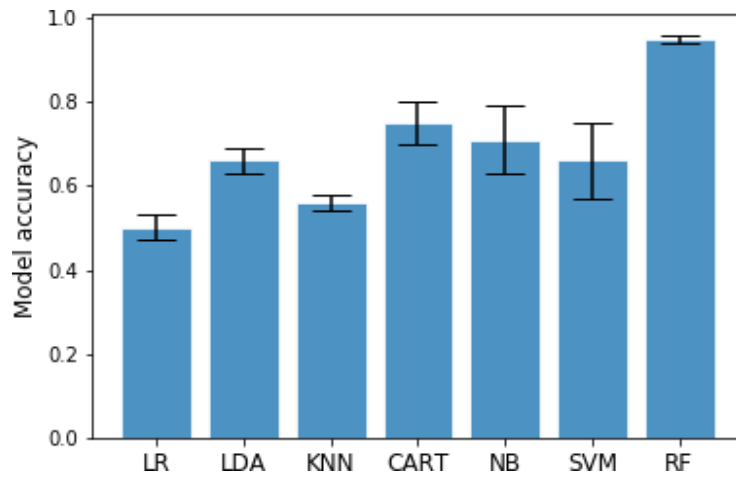


Fig. 4.4.14: Comparison of classification of ageing time on 6-month basis

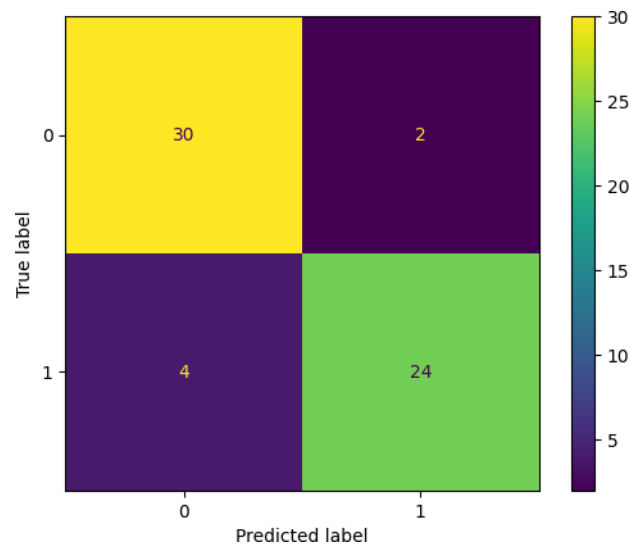


Fig. 4.4.15: Confusion matrix of ageing time on 6-month basis

4.5 Summary of Chapter IV

The findings obtained from implementing the methodologies of chapter 3 and subsequent inference are described in this chapter. The soaking kinetics gave Peleg's constant k_1 in the range 0.039 s^{-1} - 0.043 s^{-1} and k_2 in the range 0.021 - 0.033 \% db^{-1} . Midili Kucuk equation for soaking time with moisture ratio was developed for any given temperature,

$$MR = (1.009e^{-0.067t^{0.782}} + 4.54 \times 10^{-4}t)(-0.3002 + 0.025t).$$

The time (approx.) estimated for reaching 30% m.c. (wb) using Peleg's equation was found to be 111, 135, and 188 min, at 60, 50, and 40 °C respectively and using Midili Kucuk equation the time estimated were 108, 134, and 192 min at 60, 50, and 40 °C respectively. The activation energy of the soaking process was calculated to be 42.28 kJ/mol. Fick's diffusivity coefficient was found to be in the range of $2.83 - 7.92 \times 10^{-11} \text{ m}^2/\text{s}$. For, the steaming process, the first order rate constant was in the range between 0.003 to 0.001 s^{-1} with varying pressures. The time required to reach 99% gelatinization was estimated to be 21, 17, 12, 10, and 6 at 0, 0.05, 0.1, 0.15, and 0.2 MPa respectively. The activation energy of the steaming process was calculated to be $18.7 \times 10^2 \text{ kJ/mol}$. For drying kinetics, the Page model fitted the best, and the estimated time for reaching 13% m.c. (wb) was found to be 150, 165, and 185 min at 60, 50, and 40 °C respectively. The activation energy calculated for the drying process was found to be $2.49 \times 10^{-7} \text{ kJ/mol}$. The diffusion coefficient for the process of drying estimated from Crank's solution was in the range $2.60 \times 10^{-9} \text{ m}^2/\text{s}$ - $7.77 \times 10^{-9} \text{ m}^2/\text{s}$. The second objective findings include the developed ML models for process analysis of soaking, steaming, and drying using spectral data. Linear regression combined with principal components performed better in regressing moisture content during soaking with prediction score: R^2 of 0.695, and RMSE of 1.109%. PLS for the purpose of regression of NIR data with DG values worked well with an R^2 value of 0.843. The estimated highest and lowest relative percent differences (RPD) values for the test dataset were 19.79 and 0.94, respectively. For drying moisture regression with spectral, ANN performed well with cross-validation score of R^2 of 0.66 and RMSE of 7.03%. The supervised learning technique, Random Forest (RF) performed well with a prediction accuracy of 0.8952 in classifying the soaked samples based on moisture content. For the third objective, the overall composition of amylose, protein and fat contents in the rice grain remained essentially unchanged after ageing while changes were observed due to processing. However, textural and structural changes do occur. Storage at higher temperature induces more changes in pasting behavior and thermal property, as well as the rehydration property of *Komal Chaul*.

The PLS model for ageing time estimation showed regression coefficient (R^2) of 0.897. The test data performance for rehydration ratio showed a very little correlation with a low R^2 value. The prediction performance of the PLS model for gruel solid loss and cooking time showed better correlation with R^2 value of 0.77 respectively. Therefore, the classification comparison suggests that *Komal Chaul* spectra that were obtained from apparently age groups of 6 months, were showing better predictability.

4.6 References to Chapter IV

- [1] Saleh, M., Akash, M., and Ondier, G. Effects of temperature and soaking durations on the hydration kinetics of hybrid and pureline parboiled brown rice cultivars. *Journal of Food Measurement and Characterization*, 12(2):1369-1377, 2018.
- [2] Cheevitsopon, E., and Noomhorm, A. Kinetics of hydration and dimensional changes of brown rice. *Journal of Food Processing and Preservation*, 35(6):840- 849, 2011.
- [3] Wahengbam, E. D., Abdu, S., and Hazarika, M. K. Water uptake in brown rice during soaking for production of no-cooking rice. *Agricultural Engineering International: CIGR Journal*, 21(3):138-149, 2019.
- [4] Bezbaruah, B.J, and Hazarika, M. K. Generalization of temperature and thickness effects in kinetic studies of turmeric (*Curcuma longa*) slices drying. *International Food Research Journal*, 21(4):1529-1532, 2014.
- [5] Bello, M., Tolaba, M. P., Aguerre, R. J., and Suarez, C. Modeling water uptake in a cereal grain during soaking. *Journal of Food Engineering*, 97(1):95-100, 2010.
- [6] Wahengbam, E. D., Tongbram, T., and Hazarika, M. K. Drying characteristics of ready-to-eat komal chawal rice: processing and modeling. *Journal of Food Science and Technology*, 57(5):1698-1709, 2020.
- [7] Ojeda, C. A., Tolaba, M. P., and Suárez, C. Modeling starch gelatinization kinetics of milled rice flour. *Cereal Chemistry*, 77(2):145-147, 2000.
- [8] Zaroni, B., Schiraldi, A., and Simonetta, R. A naive model of starch gelatinization kinetics. *Journal of Food Engineering*, 24(1):25-33, 1995.
- [9] Chakraborty, S., Gautam, S. P., Bordoloi, T., and Hazarika, M. K. Neural network and computational fluid dynamics modeling for the gelatinization kinetics of instant controlled pressure drop treated parboiled rice. *Journal of Food Process Engineering*, 43(11): e13534, 2020.
- [10] Ahromrit, A., Ledward, D. A., and Niranjana, K. Kinetics of high pressure facilitated starch gelatinisation in Thai glutinous rice. *Journal of Food Engineering*, 79(3):834-841, 2007.
- [11] Miah, M. A. K., Haque, A., Douglass, M. P., and Clarke, B. Parboiling of rice Part I: Effect of hot soaking time on quality of milled rice. *International Journal of Food Science and Technology*, 37(5):527-537, 2002.

- [12] Yildirim, A. Kinetics and thermodynamic properties of parboiled burgos wheat (*Triticum durum*) in Turkey during drying. *Applied Ecology and Environmental Research*, 16(1):495-510, 2018.
- [13] Carvalho, T., Balbinoti, V., Nicolin, D. J., Mário, L., Jorge, D. M., Maria, R., Jorge, M., Maria, R., and Jorge, M. Parboiled rice and parboiling process. *Food Engineering Reviews*, 10:165-185, 2018.
- [14] Rao, P. S., Bal, S., and Goswami, T. K. Modelling and optimization of drying variables in thin layer drying of parboiled paddy. *Journal of Food Engineering*, 78(2):480-487, 2007.
- [15] Azzouz, S., Guizani, A., Jomaa, W., and Belghith, A. Moisture diffusivity and drying kinetic equation of convective drying of grapes. *Journal of Food engineering*, 55(4), 323-330, 2002.
- [16] Aguerre, R., Suarez, C., and Viollaz, P. E. Drying kinetics of rough rice grain. *International Journal of Food Science and Technology*, 17(6):679-686, 1982
- [17] Rizwana, S., Singamayum, F.N., Kumar, K., Pohthmi, S., and Hazarika, M. K. Drying kinetics and image-based identification of drying end point during parboiling of komal chawal. *Journal of Agriculture and Food Research*, 13:100646, 2023.
- [18] Wu, S., Wang, L., Zhou, G., Liu, C., Ji, Z., Li, Z., and Li, W. Strategies for the content determination of capsaicin and the identification of adulterated pepper powder using a hand-held near-infrared spectrometer. *Food Research International*, 163:112192, 2023.
- [19] Geng, Y., Wang, Z., Jia, L., Qin, Y., and Chen, X. Bogie fault diagnosis under variable operating conditions based on fast kurtogram and deep residual learning towards imbalanced data. *Measurement*, 166:108191, 2020.
- [20] Wang, J., Wu, X., Zheng, J., and Wu, B. Rapid identification of green tea varieties based on FT-NIR spectroscopy and LDA/QR. *Food Science and Technology (Brazil)*, 42: e73022, 2022.
- [21] Xie, L., He, X., Duan, B., Tang, S., Luo, J., Jiao, G., and Hu, P. Optimization of near-infrared reflectance model in measuring gelatinization characteristics of rice flour with a rapid viscosity analyzer (RVA) and differential scanning calorimeter (DSC). *Cereal Chemistry*, 92(5):522-528, 2015.
- [22] Sampaio, P. S., Soares, A., Castanho, A., Almeida, A. S., and Brites, C. Optimization of

- rice amylose determination by NIR-spectroscopy using PLS chemometrics algorithms. *Food Chemistry*, 242: 196-204, 2017.
- [23] Feng, Y. Z., Elmasry, G., Sun, D. W., Scannell, A. G. M., Walsh, D., and Morcy, N. Near-infrared hyperspectral imaging and partial least squares regression for rapid and reagentless determination of Enterobacteriaceae on chicken fillets. *Food Chemistry*, 138(2-3):1829-1836, 2013.
- [24] Brasil, Y. L., Cruz-Tirado, J. P., and Barbin, D. F. Fast online estimation of quail eggs freshness using portable NIR spectrometer and machine learning. *Food Control*, 131:108418, 2022.
- [25] Cruz-Tirado, J. P., Lucimar da Silva Medeiros, M., and Barbin, D. F. On-line monitoring of egg freshness using a portable NIR spectrometer in tandem with machine learning. *Journal of Food Engineering*, 306:110643, 2021.
- [26] Nolasco-Perez, I. M., Rocco, L. A. C. M., Cruz-Tirado, J. P., Pollonio, M. A. R., Barbon, S., Barbon, A. P. A. C., and Barbin, D. F. Comparison of rapid techniques for classification of ground meat. *Biosystems Engineering*, 183:151-159, 2019.
- [27] Oliveira, M. M., Cruz-Tirado, J. P., Roque, J. V., Teófilo, R. F., and Barbin, D. F. Portable near-infrared spectroscopy for rapid authentication of adulterated paprika powder. *Journal of Food Composition and Analysis*, 87:103403, 2020.
- [28] Park, S., Myongkyoon, Y., Dong Gyun, Y., Cheorun, J., and Ghiseok, K. VIS/NIR hyperspectral imaging with artificial neural networks to evaluate the content of thiobarbituric acid reactive substances in beef muscle. *Journal of Food Engineering*, 350: e111500, 2023.
- [29] María J, and Martelo-Vidal Vázquez, M. Application of artificial neural networks coupled to UV–VIS–NIR spectroscopy for the rapid quantification of wine compounds in aqueous mixtures. *CyTA-Journal of Food*, 13(1):32-39, 2015.
- [30] Li, X., and He, Y. Discriminating varieties of tea plant based on Vis/NIR spectral characteristics and using artificial neural networks. *Biosystems Engineering*, 99(3):313-321, 2008.
- [31] Kamruzzaman, M., Kalita, D., Ahmed, M. T., G, E., and Makino, Y. Effect of variable selection algorithms on model performance for predicting moisture content in biological materials using spectral data. *Analytica chimica acta*, 1202:339390, 2022.

- [32] Chrastil, J., and Zarins, Z. M. Influence of Storage on Peptide Subunit Composition of Rice Oryzenin. *Journal of Agricultural and Food Chemistry*, 40(6):927-930, 1992.
- [33] Zhou, Z., Wang, X., Si, X., Blanchard, C., and Strappe, P. The ageing mechanism of stored rice: A concept model from the past to the present. *Journal of Stored Products Research*, 64:80-87, 2015.
- [34] Sowbhagya, C. M., and Bhattacharya, K. R. Changes in pasting behaviour of rice during ageing. *Journal of Cereal Science*, 34(2):115-124, 2001.
- [35] Zhou, Z., Robards, K., Helliwell, S., and Blanchard, C. Ageing of Stored Rice: Changes in Chemical and Physical Attributes. *Journal of Cereal Science*, 35(1):65-78, 2002.
- [36] Wahengbam, E. D., Green, B. D., and Hazarika, M. K. Characterization of a novel folic acid-fortified ready-to-eat parboiled rice. *Cereal Chemistry*, 96(3):439-446, 2019.
- [37] Dutta, H., and Mahanta, C. L. Laboratory process development and physicochemical characterization of a low amylose and hydrothermally treated ready-to-eat rice product requiring no cooking. *Food and Bioprocess Technology*, 7(1):212-223, 2014.
- [38] Wahengbam, E. D., and Hazarika, M. K. Quality of ready-to-eat komal chawal produced by brown rice parboiling method. *Journal of Food Science and Technology*, 56(1):187-199, 2019.
- [39] Kumari, S., Chakraborty, S., Choudhary, Kumar, A., Boiragi, A., Das, O., and Hazarika, M. K. Neuro-fuzzy interface and mathematical modeling of rehydration kinetics and dynamic vapor sorption behavior of novel no-cooking rice. *Journal of Food Process Engineering*, 46(4): e14299, 2023.
- [40] Zhou, Z., Robards, K., Helliwell, S., and Blanchard, C. Effect of storage temperature on cooking behaviour of rice. *Food Chemistry*, 105(2):491-497, 2007.
- [41] Huang, Y. C., and Lai, H. M. Characteristics of the starch fine structure and pasting properties of waxy rice during storage. *Food Chemistry*, 152:432-439, 2014
- [42] Saikrishna, A., Dutta, S., Subramanian, V., Moses, J. A., and Anandharamakrishnan, C. Ageing of rice: A review. *Journal of Cereal Science*, 81:161-170, 2018.
- [43] Ayoub, A., Ohtani, T., and Sugiyama, S. Atomic force microscopy investigation of disorder process on rice starch granule surface. *Starch/Stärke*, 58(9):475-479, 2006.
- [44] Mehmood, T., Liland, K. H., Snipen, L., and Sæbø, S. A review of variable selection

- methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 118:62-69, 2012.
- [45] Miao, X., Miao, Y., Tao, S., Liu, D., Chen, Z., Wang, J., Huang, W. and Yu, Y., Classification of rice based on storage time by using near infrared spectroscopy and chemometric methods. *Microchemical Journal*, 171:106841, 2021.
- [46] Xue, J. H., and Titterington, D. M. Comment on discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Neural Processing Letters*, 28(3):169-187, 2008.