Chemometric classification of geothermal and non-geothermal ethanol leaf extract of seurapoh \( (Chromolaena odorata \text{ Linn}) \) using infrared spectroscopy

A Abubakar\(^1,2,7\), H Yusuf\(^3\), M Syukri\(^4\), R Nasution\(^5\), T Karma\(^5\), A A Munawar\(^6,7\) and R Idroes\(^5,8\)

\(^1\) Graduate School of Mathematics and Applied Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
\(^2\) Gastroenterology Division, Internal Medicine Department, School of Medicine, Universitas Syiah Kuala, dr. Zainoel Abidin Teaching Hospital, Banda Aceh 23111, Indonesia
\(^3\) Pharmacology Department, School of Medicine, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
\(^4\) Nephrology and Hypertension Division, Internal Medicine Department, School of Medicine, Universitas Syiah Kuala, dr. Zainoel Abidin Teaching Hospital, Banda Aceh 23111, Indonesia
\(^5\) Department of Chemistry, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
\(^6\) Atsiri Research Centre (ARC) PUI Nilam Aceh, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
\(^7\) Agricultural Technology Department, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
\(^8\) Department of Pharmacy, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia

E-mail: azzakiabubakar@yahoo.com

Abstract. This work reports on the chemometric classification of ethanol leaf extract of seurapoh \( (Chromolaena odorata \text{ Linn}) \) obtained from geothermal and non-geothermal area. This analysis aims to identify the correlation between the ethanol leaf extract of seurapoh (ELES) collected from three locations, one geothermal area (Ie Seuum) and two non-geothermal areas (Lhoknga Beach, Samahani), Aceh Besar Province, Indonesia. The extraction was made by ethanol 96\% using maceration. After extraction, each sample's infrared spectrum was acquired using near-infrared (NIR) spectroscopy at a wavenumber range of 10,000 cm\(^{-1}\) – 4,000 cm\(^{-1}\). The spectral data were preprocessed with the MSC method to remove the light scattering effect during spectral acquisition. The classification of ELES was conducted using Principle Component Analysis (PCA), yielding plot scores as follows, PC1 and PC2 were capable of explaining 97\% of the total data (PC1 = 87\%, PC2 = 10\%). The results also revealed very well discrimination of the samples based on their respective sampling locations. It further implied that the ELES had significant different NIR spectral profiles, which were following their sampling locations. Hence, it can be concluded that there is a possibility of different compounds contained within the ELES, collected from the geothermal area and two other non-geothermal areas.
1. Introduction
Seurapoh is a name given by the local community in Aceh, Indonesia, for the Kirinyuh plant (*Chromolaena odorata* Linn), a family Asteraceae species. Acehnese people, traditionally, have observed the application of seurapoh leaf in the wound healing treatment. Based on the previous reports, a seurapoh leaf contains several significant compounds such as tannins, phenols, flavonoids, saponins, and steroids. As an addition to the compounds as mentioned earlier, it also contains essential oils, including α pinene, cadinene, camphora, limonene, β-caryophyllene, and candinol isomer [1].

The quality of a medicinal plant is determined by its secondary metabolites [2]. Secondary metabolites are the results of the interaction between plants and their environment. Rather than on its primary metabolite, the correlation between a plant and its environment is more likely to affect the secondary metabolite content [3]. A secondary metabolite content of a plant is strongly affected by the surrounding environment because each location has a different characteristic for one another. Environmental factors (soil, water, and climate) contribute significantly to the formation of secondary metabolites of a plant [3]. The same type of plant, from the same species, grows under different environmental conditions, was reported to have direct differences in producing and accumulating secondary metabolites [4]–[6]. Furthermore, nowadays, the metabolism profile has become an essential tool in understanding an organism's systematic responses against the changing environmental condition [7]. One of the locations with unique characteristics, in comparison to other locations, is geothermal manifestation area [8]–[10].

Seurapoh is an easily found weed plant area with a tropical climate (such as Indonesia). Based on the observation, seurapoh plant was found in coastal and geothermal manifestation (such as Ie Seuum) areas. As reported previously, Ie Seuum is a chloride-dominant geothermal manifestation [11]. Therefore, it is interesting to identify the difference of seurapoh leaf collected from geothermal and non-geothermal areas. For initial identification, one of the usable methods is a combination of infrared spectroscopy and chemometric.

Several analytical instruments frequently used in an organic samples, namely Mass Spectrometry [12–15], image processing [16], laser [17], [18], Uv-Vis [19]–[21], and Fourier Transform Infrared (FTIR) [22]. FTIR is the most suitable one for this purpose. The use of FTIR as an identification method is based on the evidence that the infrared spectra are a fingerprint pattern [23], indicating the unique composition characteristics of a sample [24]. Owing to the complexity of IR spectra, direct visual interpretation can be challenging. For that reason, a multivariate analytical technique is proposed to assist the interpretation [25]–[28]. A multivariate statistical method, which can be used to identify the infrared spectra, is Principal Component Analysis (PCA) [29], [30]. PCA's main objective is to maintain the majority of data variance within a complex data set by reducing the dimension of major intercorrelated variables into a single main component [31]. Based on above description, this work conducted a classification of seurapoh leaf, collected from different locations to identify the discrimination pattern using a combination of infrared spectroscopy and chemometric.

2. Research methodology
2.1. Tools and materials
Tools used in this research included a set of maceration equipment, rotary evaporator (RE), NIR spectrometer, and computer set (for data processing). Meanwhile, the materials included seurapoh leaf (collected from Ie Seuum, Lhoknga Beach, and Samahani) and analytical grade ethanol 96% (purchased from Merck KgaA, Germany).

2.2. Research procedure
2.2.1. Sample collection. Samples are the leaves of seurapoh plants (*Chromolaena odorata* Linn) which were collected from 3 locations, namely geothermal manifestation area Ie Seuum, Coastal area Lhoknga, and Samahani area, where 3 sampling points were predetermined with the purposive method in each sampling location.
2.2.2. Sample preparation. After collected from the sampling locations, the samples were washed and air-dried for a week. It is then grounded to obtain the powder, where the powder was used for extraction using ethanol 96%. The maceration lasted for 48 h, with occasional stirring. Afterward, the sample was filtered, and the filtrate was evaporated using RE to remove the solvent. This procedure yielded ethanol leaf extract of seurapoh.

2.2.3. Spectrum acquisition using NIR spectroscopy. Seurapoh leaf extract, collected from each location, was then analyzed using the NIR spectrometer to acquire the spectral profile at a wavenumber range of 10,000 – 4,000 cm\(^{-1}\) (resolution 4 cm\(^{-1}\)). Therefore, in total, there were 9 spectra obtained from the analysis.

2.2.4. Improving NIR spectra. Before analyzed, the spectral data from NIR analysis was priorly pretreated/preprocessed. Data preprocessing is an essential step in the chemometric analysis, either in PCA or other chemometric methods [32]. The main aim of this procedure is to generate the fittest spectral data to obtain an optimal result. By removing the uncorrelated variance from the samples' information, the analysis can be more focused on the observed variance. In this research, data preprocessing was conducted using the Multiplicative Scatter Correction (MSC) method.

2.2.5. Classification of ethanol leaf extract of seurapoh with PCA method. PCA was conducted on the data which had been priorly preprocessed. The analysis was conducted using software assistance, namely The Unscrambler 10.4. The discrimination of ethanol leaf extract of seurapoh using PCA was represented by the value of two-dimensional Principle Component (PC). This plot gave information about the pattern within the leaf extract samples of seurapoh. Usually, two initial PC plots gave the most data variance.

3. Results and discussion
In this research, samples used were ethanol leaf extract of seurapoh, collected from several different locations, including geothermal manifestation Le Seeum, Lhoknga, and Samahani. All the three sampling locations were in Aceh Besar Regency, where three sampling points were proposed in each sampling locations, giving a total of 9 sampling points. The collected samples were washed and dried for a week, followed by grounding them into powder. It is then macerated using ethanol 96% for 48 h. Afterward, the samples were filtered, and the filtrate was collected, followed by an evaporation step of the filtrate using RE to obtain the concentrated extract. Samples were then analyzed with a NIR spectrometer for the acquisition of infrared spectra.

3.1. NIR spectra of seurapoh leaf extract
NIR spectra acquisition of the simplisia and leaf extract of seurapoh was conducted at a wavenumber range of 10,000 – 4,000 cm\(^{-1}\). The spectral image of the ethanol leaf extract of seurapoh is presented in Figure 1. Based on Figure 1, several absorbance bands are observable and attributed to specific moieties. The absorbance band at wavenumber 8,500 – 9,000 cm\(^{-1}\) is assigned to C-H and C-C aromatic moieties. The presence of CH, CH\(_3\), CH\(_2\), and OH is indicated by the absorbance band at 5,800 – 5,900 cm\(^{-1}\). Meanwhile, an absorbance band at 5,300– 5,500 cm\(^{-1}\) is associated with -OH stretching. The vibration of -OH ethanol and sugar are observed at 4,800 – 5,000 cm\(^{-1}\). Absorbance bands at a range of 4,440 – 4,500 cm\(^{-1}\) indicate the presence of CH and OH moieties [33].
3.2. NIR spectral data preprocessing

To optimize the chemometric analysis, preprocessing is crucial should be taken into account. Preprocessing was conducted on the obtained IR spectral data to avoid the overlapping between the observational target data with non-target data (could be generated from matrix effect including light scattering and noise signal effects) [34]. Data obtained from the analysis are highly expected to contain non-target data from the light scattering effect. However, it is an exception for aqueous or liquid samples that are entirely dissolved, allowing the particle within the sample to appear in the same size with even distribution. A solid sample can be confident that the light scattering effect always occurs due to the roughness of the sample surface and uneven particle sizes. In this work, the data were preprocessed using the Multiplicative Scatter Correction (MSC) method. It aims to reduce the light scattering effect in the NIR spectra by rotating each spectrum to find the best similarity with the standardized spectrum, possibly becoming the mean spectrum [35].

Figure 1. NIR spectrum of ethanol leaf extract of seurapoh.

Figure 2. NIR spectra of seurapoh leaf extract after preprocessed using MSC.
3.3. Classification of seurapoh leaf extract with PCA

PCA is a multivariate statistical method often used to generate a discriminant model for closely related plants and unobserved pattern recognition techniques [36]. PCA is used to reduce the data and seek information for variable combinations, explaining a significant trend within a set of data [37]. Classification of ethanol leaf extract of seurapoh using PCA was represented by two-dimensional PC value. The plot informs the pattern generated by the NIR-spectra of seurapoh leaf. The plot for the two initial PC usually gives the most data variance. Principally, the PCA plot can be interpreted, where the closer a sample to one another, the closer their similarities are.

Figure 3 exhibits that PC1 and PC2 were able to explain the total variance of up to 97% (PC1= 87%, PC2=10%). The pattern of the classification of seurapoh leaf extract, shown on the PCA plot, can distinguish the ethanol leaf extracts of seurapoh, collected from 3 different locations. Based on the plot, it revealed that the ethanol leaf extract of seurapoh, collected from the three locations, possess different chemical compounds. According to [38-40], this finding is where the content of secondary metabolites in a plant is positively affected by the surrounding environment. It is because each location has a different characteristic for one another. Environmental factors (soil, water, and climate) significantly contribute to forming the secondary metabolite of a plant [3]. The same type of a plant from the same species, growing under different environmental conditions, had been observed to produce and accumulate secondary metabolites differently [4]. Moreover, [8] reported that the geothermal manifestation Ie Seuum had dominant chloride content than other mineral salts. Other than affected by the soil mineral, the geothermal-induced temperature in the manifestation area (Ie Seuum) also theoretically affected the secondary metabolite content within the seurapoh leaf. Although the samples taken from Lhoknga and Samahani came from non-geothermal areas, based on the PCA plot the samples from the two locations were also separated. This was due to different environmental conditions in Lhoknga and Samahani, where Lhoknga is a coastal area, and Samahani is a non-geothermal area that is far from the coast and also far from the geothermal area. This allows differences in soil elements and environmental conditions that cause differences in chemical content. Based on the analysis that has been done, it shows that the combination method of NIR-PCA can work optimally to identify seurapoh leaves from different locations.

![PCA plot score of ethanol leaf extract of seurapoh.](image)
4. Conclusion

Based on the results obtained in this work, it can be concluded that the secondary metabolite content of seurapoh plant is very much affected by the environmental condition, confirmed by the PCA plot. It can be observed that the samples fell under the same classification group, according to the respective sampling location. It suggests the presence of a significant difference in the NIR spectral profile among the ethanol leaf extracts of seurapoh, collected from the geothermal manifestation area (Ie Seuum) and non-geothermal area (Lhoknga Beach and Samahani). Further research can be focused to observe the difference in secondary metabolite content, based on the sampling location.

References

[1] Inya-agha S I, Oguntimein B O, Sofowora A and Benjamin T V 1987 *Int. J. Crude Drug Res.* **25** 49–52
[2] Dong J, Ma X, Wei Q, Peng S and Zhang S 2011 *Ind. Crops Prod.* **34** 1607–14
[3] Liu W, Wang D, Hou X, Yang Y, Xue X, Jia Q, Zhang L, Zhao W and Yin D 2018 *Chem. Biodivers.* **15** e1800114
[4] Sampaio B L, Edrada-Ebel R and Da Costa F B 2016 *Sci. Rep.* **6** 29265
[5] Nuraskin C, Marлина, Idroes R, Soraya C and Djufri 2020 *Rasayan J. Chem.* **15** e1800114
[6] Sampaio B L, Edrada-Ebel R and Da Costa F B 2016 *Sci. Rep.* **6** 29265
[7] Westen L A, Skoneczny D, Weston P A and Weidenhamer J D 2015 *J. Allelochem. Interact.* **1** 15–27
[8] Idroes R, Yusuf M, Alatas M, Subhan, Lala A, Muslem, Suhendra R, Idroes G M, Suhendrayatna, Marvan and Riza M 2019 *IOP Conf. Ser. Mater. Sci. Eng.* **523** 012010
[9] Idroes R, Yusuf M, Alatas M, Subhan, Lala A, Muhammad, Suhendra R, Idroes G M and Marwan 2019 *IOP Conf. Ser. Mater. Sci. Eng.* **523** 012012
[10] Idroes R, Yusuf M, Saiful S, Alatas M, Subhan, Lala A, Muslem M, Suhendra R, Idroes G M, Marwan M and Mahlia T M I 2019 *Energies* **12** 4442
[11] Idroes R, Yusuf M, Alatas M, Subhan, Lala A, Saiful, Suhendra R, Idroes G M and Marwan 2018 *IOP Conf. Ser. Mater. Sci. Eng.* **334** 012002
[12] Estevam E, Griffins S, Nasim J, Zieliński D, Aszyk J, Osowicka M, Dawidowska N, Idroes R, Bartoszek A and Jacob C 2015 *Nat. Prod. Commun.* **10** 1733–8
[13] Prativi S U T, Lagendijk E L, de Weert S, Idroes R, Hertiani T and den Hondel C Van 2015 *Int. J. Appl. Res. Nat. Prod.* **8** 1–13
[14] Earlia N, Muslem, Suhendra R, Amin M, Prakoeswa C R S, Khairan and Idroes R 2019 *Sci. World J.* **2019** 1–7
[15] Earlia N, Rahmad R, Amin M, Prakoeswa C, Khairan K and Idroes R 2019 *Sains Malaysiana* **48** 1019–24
[16] Suhartono E, Setiawan B, Santosa P B, Idroes R and Indrawan M S 2019 *Journal of Physics: Conference Series* vol 1374 (IOP Publishing) p 12057
[17] Hedwig R, Lahna K, Idroes R, Karnadi I, Tanra I, Iqbal J, Kwaria D, Kurniawan D P, Kurniawan K H, Tjia M O and Kagawa K 2019 *Microchem. J.* **147** 356–64
[18] Lahna K, Idroes R, Idris N, Abdulmadjid S N, Kurniawan K H, Tjia M O, Pardede M and Kagawa K 2016 *Appl. Opt.* **55** 1731
[19] Suhartono E, Thalib I, Aflanie I, Noor Z and Idroes R 2018 *IOP Conf. Ser. Mater. Sci. Eng.* **350** 12008
[20] Hasanah U, Sani N D M, Heng L Y, Idroes R and Safitri E 2019 *Biosens. Sens.* **9** 135
[21] Hasanah U, Setyowati M, Efendi R, Muslem M, Md Sani N D, Safitri E, Yook Heng L and Idroes R 2019 *Biosens. Sens.* **9** 60
[22] Suhartono E, Noor Z, Edyson, Budianto W Y and Idroes R 2019 *AIP Conf. Proc.* **2108** 020025
[23] Hasanah U, Setyowati M, Edwarsyah, Efendi R, Safitri E, Idroes R, Heng L Y and Sani N D 2018 *IOP Conf. Ser. Mater. Sci. Eng.* **523** 12014
[24] Sanchez P M, Pauli E D, Scheel G L, Rakosевич M, Bruns R E and Scarminio I S 2018 *J. Braz.
[25] Idroes R, Husna I, Muslem, Mahmudi, Rusyana A, Helwani Z, Idroes G M, Suhendra R, Yandri E, Rahimah S and Sasmita N R 2019 IOP Conference Series: Earth and Environmental Science vol 364 (IOP Publishing) p 12020

[26] Idroes R, Noviandy T R, Maulana A, Suhendra R, Sasmita N R, Muslem M, Idroes G M and Irvanizam I 2019 Algorithm Int. Rev. Model. Simulations 12 373

[27] Husna I, Rusyana A, Muslem, Idroes G M, Suhendra and R and Idroes R 2020 IOP Conf. Ser. Mater. Sci. Eng. 796 012064

[28] Idroes R, Japnur A F, Suhendra R and Rusyana A 2019 IOP Conf. Ser. Mater. Sci. Eng. 523 12007

[29] Wahidah S, Khairi, Lelifajri, Idroes R, Rahmadi, Lala A, Mahmudi, Muslem and Japnur A F 2019 IOP Conf. Ser. Earth Environ. Sci. vol 364 (IOP Publishing) p 12040

[31] Kumar R and Sharma V 2018 TrAC Trends Anal. Chem. 105 191–201

[32] Brereton R G, Jansen J, Lopes J, Marini F, Pomerantsev A, Rodionova O, Roger J M, Walczak B and Tauler R 2017 Anal. Bioanal. Chem. 409 5891–9

[33] Schwanninger M, Rodrigues J C and Fackler K 2011 J. Near Infrared Spectrosc. 19 287–308

[34] Rinnan Å 2014 Anal. Methods 6 7124–9

[35] Ramadhan S, Munawar A A and Nurba D 2016 J. Ilm. Mhs. Pertan. 1 954–60

[36] Purwakusumah E D, Rafi M, Safitri U D, Nurcholis W and Zaim Adzkiya M A 2014 Agritech J. Fak. Teknol. Pertan. UGM 34 82–7

[37] Gad H A, El-Ahmady S H, Abou-Shoer M I and Al-Azizi M M 2013 Phytochem. Anal. 24 1–24

[38] Guo L, Wu Y, Liu M, Ge Y and Chen Y 2018 J. Sci. Food Agric. 98 3057–65

[39] Munawar AA, Yunus Y, Devianti, Satriyo P 2020 Data Br. 30.

[40] Devianti D, Sufardi S, Zulfahrizal Z, Munawar AA 2019 In: IOP Conference Series: Materials Science and Engineering. Institute of Physics Publishing.