

## Pengolahan Data Arus Laut menggunakan Bahasa Program R

Diah Pitaloka, Onie Wiwid Jayanthi\*, Ary Giri Dwi Kartika, Ashari Wicaksono,  
Muh. Syaifullah, Irsyadatul Fikriah

Program Studi Ilmu Kelautan, Jurusan Kelautan dan Perikanan, Fakultas Pertanian, Universitas Trunojoyo Madura  
Jl. Raya Telang, Kamal, Bangkalan, Jawa Timur 69162 Indonesia  
Email: onie.jayanti@trunojoyo.ac.id

### Abstrak

Penggunaan bahasa pemrograman merupakan salah satu dari penerapan kemudahan teknologi, melalui pemrosesan, pengolahan, serta analisis data, berupa *output* data yang jelas dan akurat secara cepat dan relevan. Bahasa program R dinilai cukup baik sebagai alat simulasi model, baik melalui kemampuan kalkulasi dan manipulasi data, maupun pada kemampuan peragaan grafik dan visualisasi data. Tujuan dari penelitian ini ialah mengetahui tahapan pengolahan data arus menggunakan bahasa program R. Keunggulan bahasa pemrograman tersebut ialah bersifat *open source* dan dapat digunakan untuk mengolah *big data*. Pengambilan data arus menggunakan instrumen *Acoustic Doppler Current Profiler* (ADCP) dilakukan secara insitu menggunakan metode eulerian. Perekaman data dilakukan selama 3 hari di kawasan Perairan Pademawu, Pamekasan, Madura. Data diolah menggunakan bahasa pemrograman R versi 4.1.2. dengan memanfaatkan *software* R-Studio. Pengolahan data tersebut akan menghasilkan grafik dan *stick plot* arah dan kecepatan komponen arus U (Timur-Barat) dan V (Selatan-Utara). Visualisasi tersebut menginterpretasikan karakteristik arus di lokasi penelitian. Kesimpulan dari penelitian ini ialah tahapan pengolahan data arus menggunakan bahasa program R diantaranya yaitu pre-pengolahan data, pengolahan data, dan post-pengolahan data. Pada tahapan post-processing data, dihasilkan 442 *records* dan 21 *variable dataset* sebagai bahan visualisasi *stick plot* dan grafik garis U-V yang dapat menginterpretasikan karakteristik arus laut di kawasan penelitian.

**Kata kunci :** Program R, Arus, ADCP, Pademawu

### Abstract

#### *Ocean Current Data Processing using the R Program*

*Using a programming language is one of the applications of technological convenience, through data processing, processing, and analysis, in the form of clear and accurate data output that is fast and relevant. The R programming language is considered quite good as a model simulation tool, both through its calculation and data manipulation capabilities, as well as in its graphical display capabilities and data visualization. The aim of this research is to know the stages of data processing using the R programming language. The advantage of this programming language is that it is open-source and can be used to process big data. Current data collection using the Acoustic Doppler Current Profiler (ADCP) instrument was carried out in situ using the Eulerian method. Data recording was carried out for 3 days in the Pademawu Waters, Pamekasan, Madura. Data is processed using the R programming language version 4.1.2. it is using R-Studio software. Processing the data will produce graphs and stick plots of the direction and velocity of the U (East-West) and V (South-North) current components. The visualization interprets the characteristics of the currents at the research location. This study concludes that the stages of processing data flow using the R programming language include data pre-processing, data processing, and post-data processing. At the data post-processing stage, 442 records and 21 dataset variables were produced as visualization materials for stick plots and U - V line graphs that can interpret the characteristics of ocean currents in the research area.*

**Keywords :** Program R, Current, ADCP, Pademawu

## PENDAHULUAN

Perkembangan teknologi memiliki grafik yang sejajar dengan perkembangan zaman secara global. Hal ini sejalan dengan masuknya era industri 4.0, yang merupakan transformasi manufaktur mesin ke era manufaktur digital. Teknologi menuntun manusia pada hidup yang lebih mudah dan lebih praktis dalam menjalankan segala sesuatu (Thimbleby, 2013 dst.). Kemudahan teknologi tersebut menjadi sebuah kebutuhan dan terus dikembangkan di berbagai bidang kehidupan dan profesi, termasuk dikalangan para peneliti, *data scientist*, maupun *data analyst*. Para *data analyst* mengaplikasikan hal tersebut pada pemrosesan, pengolahan, serta analisis data, dengan menghasilkan *output* data yang jelas dan akurat, secara cepat dan relevan (Rasheva-Yordanova *et al.*, 2018 dst.). Berdasarkan hal tersebut penggunaan bahasa pemrograman menjadi salah satu penerapan penggunaan kemudahan teknologi (Siregar dan Nasution, 2020).

Penggunaan bahasa pemrograman dinilai dapat menyederhanakan data secara terstruktur dalam penerapan pemrosesan data, sehingga menjadi lebih mudah untuk dipahami dan dianalisis. Metode statistika deskriptif diterapkan dalam hal tersebut (Yuniarti, 2022). Oleh karena itu bahasa pemrograman menjadi *tools* yang cukup populer dikalangan para *data analyst*. Selain beragam sisi fungsional dalam pengolahan data, kemudahan pengembangan program serta kemudahan dalam mengakses *software* juga cukup diminati. Oleh karena itu penggunaan *software* dengan kategori *open source* dapat menjadi alternatif yang dipertimbangkan (Budiharto dan Rachmawati, 2013).

Kaya *et al.* (2019) mendefinisikan *open source software* sebagai *software* yang dirilis dan didistribusikan secara terbuka dan tersedia bagi siapa saja. Salah satu jenis dari *open-source software* tersebut ialah bahasa program R. Bahasa R umumnya diaplikasikan dalam bidang statistika untuk pengolahan serta komputasi data. Bahasa program ini dapat diakses secara bebas dan intensif tanpa harus khawatir terhadap lisensi yang berbayar. Bahasa program R juga dapat menampilkan visualisasi berupa grafik melalui *packages*, yang dapat diaplikasikan di bidang penelitian serta industri (Yanti, 2021). Kemampuan bahasa program R yang dapat memadukan serta memanipulasi data, kalkulasi serta peragaan grafik, cukup baik dan memadai sebagai *tool* simulasi model yang membutuhkan

banyak peragaan grafik dan visualisasi data, dengan resolusi yang dapat diatur secara penuh oleh pengguna (Yudistira, 2013). Potensi tersebut dapat diterapkan sebagai alat interpretasi suatu instrumen dan analisis yang berguna di dunia penelitian, salah satunya di bidang oseanografi dan observasi kelautan. Keunggulan lain yang dapat diperoleh dari penggunaan bahasa program R ialah penampungan dan pengolahan data dalam skala besar.

Data observasi kelautan merupakan jenis data yang besar karena memiliki karakteristik 3V, yaitu *volume*, *velocity*, dan *variety*. Oleh karena itu, data yang dihasilkan dari instrumen pengukuran observasi kelautan juga memiliki volume yang besar. Salah satu dari instrumen pengukuran tersebut ialah *Acoustic Doppler Current Profiler* (ADCP) (Liu *et al.*, 2017). ADCP mampu melakukan pengukuran arus dalam jangka waktu yang relatif panjang, pada kedalaman 30-35 meter dibawah laut di beberapa lapis kedalaman air secara vertikal. Kemampuan pengukuran dengan interval yang cukup kecil ( $\pm 3$  detik) dengan resolusi tinggi juga menjadi kelebihan instrumen tersebut (Indrayanti *et al.*, 2020). ADCP dapat menghasilkan *dataset* pengukuran arus dalam ukuran yang cukup besar, terutama jika dilakukan perekaman arus dalam kurun waktu yang panjang. Hal tersebut menjadi mustahil jika pengolahan data dilakukan menggunakan *software* pengolahan data statistik yang kurang praktis dalam menampung dan mengolah *big data*. Dalam hal ini, penggunaan bahasa pemrograman dapat menjadi pilihan yang tepat dalam pengolahan data arus ADCP (Cholissodin dan Riyandani 2016).

Arus merupakan salah satu dari parameter hidro-oseanografi perairan, dan dapat menjadi langkah awal dalam monitoring kondisi perairan. Arus laut merupakan perpindahan masa air laut yang dapat mengendalikan suatu transport partikel dan objek di suatu perairan, dimana komponen kecepatan dan arah arus menjadi penting untuk dipelajari dalam mengetahui informasi jangkauan perpindahan objek tersebut (Kumagai *et al.*, 2018). Peran arus juga dipaparkan oleh Hu dan Jun (2014) dalam penelitiannya, bahwa kondisi atmosfer dapat terdampak oleh pergerakan arus, melalui transport suhu lautan dan samudera yang dikendalikan arus tersebut. Monitoring ataupun penelitian mengenai arus dapat menjadi suatu hal yang rumit. Terkait hal tersebut pengamatan pola pergerakan dan kecepatan arus dapat menjadi suatu alternatif yang mudah melalui pendekatan matematik pemodelan

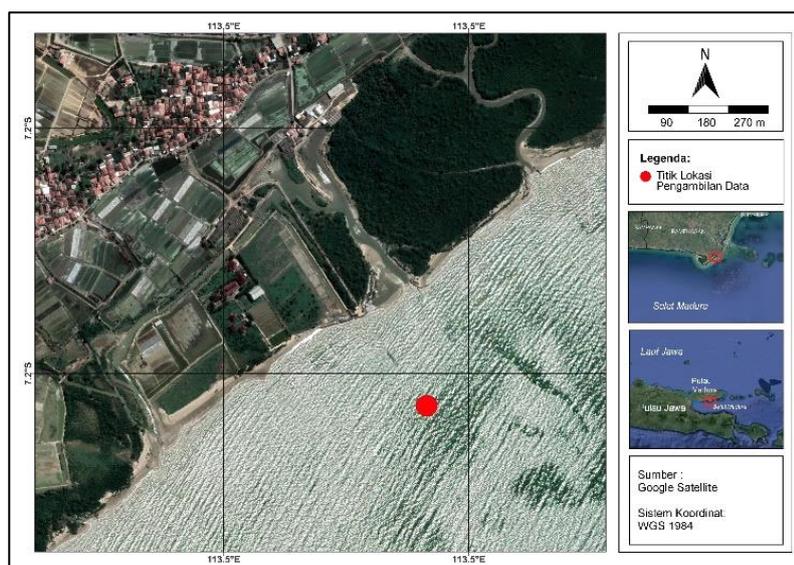
fenomena alam. Pemodelan tersebut dapat difungsikan untuk memperoleh informasi data, baik di masa sekarang maupun prediksi di masa yang akan datang (Wisha *et al.*, 2015). Penelitian serupa yang berkaitan dengan pemodelan 2D terhadap pola arus dan menghasilkan output berupa karakteristiknya juga dilakukan oleh beberapa penelitian terdahulu diantaranya oleh Dewi *et al.* (2021) yang mengkaji beberapa faktor hidro-oseanografi di wilayah ekosistem mangrove, di Desa Pagatan Besar, Tanto *et al.* (2017) yang mengkaji di wilayah Perairan Teluk Benoa, Bali, Naulita (2016) yang melakukan kajian di Kanal Labani, Selat Makassar, dan Wattimena *et al.* (2014) yang mengkaji pola dan sebaran arus di Perairan Halmahera. Sedangkan untuk penelitian terdahulu terkait pengolahan data yang menampilkan visualisasi menggunakan bahasa R diantaranya Lemenkova (2018), yang melakukan analisis statistika geomorfologi Palung Mariana, dan Gatto *et al.* (2015), yang melakukan visualisasi terhadap data proteomik.

Tujuan dari penelitian ini ialah mengetahui tahapan pengolahan data arus menggunakan bahasa program R. Pengolahan data tersebut akan menghasilkan grafik garis (*line chart*) dan *stick plot* yang dihasilkan sebagai visualisasi data arus U (Timur-Barat) dan V (Selatan-Utara) secara horizontal. Visualisasi tersebut dapat memudahkan analisis dan interpretasi karakteristik arus di kawasan Perairan Pademawu yang dipilih sebagai lokasi penelitian. Perairan Pademawu menyediakan sumber air baku pada tambak garam

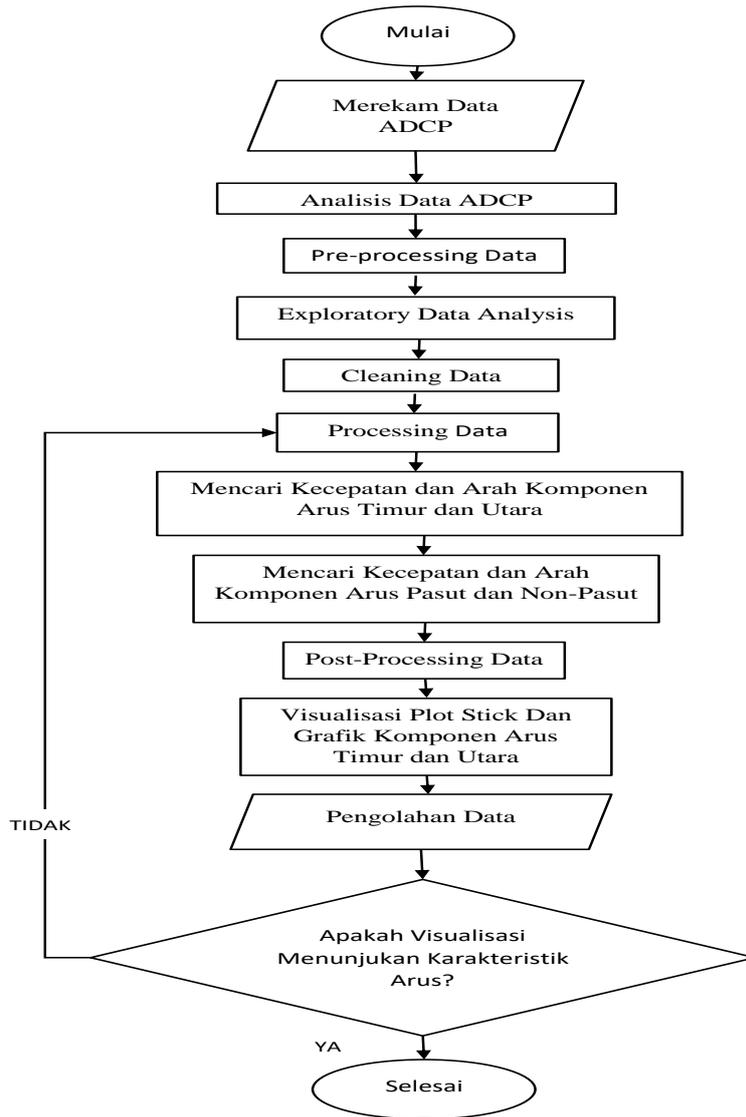
yang terletak di pesisir Desa Pademawu, Kabupaten Pamekasan, Madura. Studi hidro-oseanografi pada lokasi tersebut perlu dilakukan dengan harapan dapat menjadi acuan pengelolaan sumber air baku tambak garam dan memberikan manfaat bagi masyarakat pesisir Desa Pademawu.

## MATERI DAN METODE

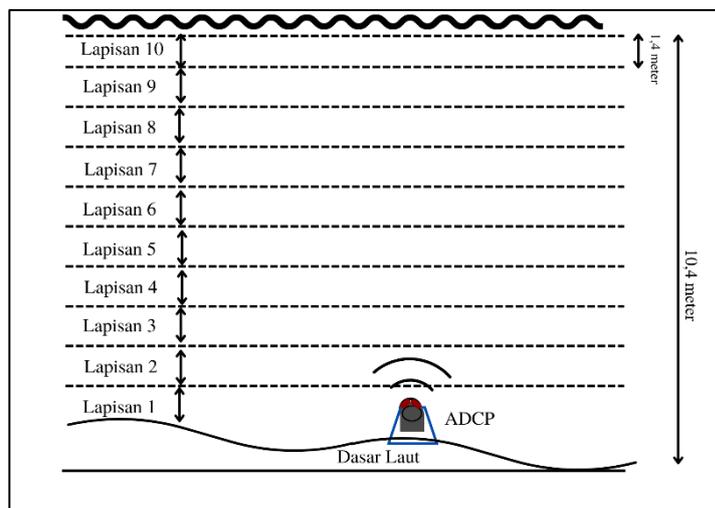
Lokasi penelitian berada di Perairan Pademawu, Kabupaten Pamekasan, Madura. Data primer penelitian merupakan data yang didapatkan melalui pengukuran arus yang dilakukan selama 3 hari, pada tanggal 13 - 15 November 2020. Instrumen yang digunakan ialah *Acoustic Doppler Current Profiler* (ADCP) AWAC AST 1MHz w/48v. Pengukuran dilakukan dengan *men-deploy* ADCP pada satu titik menggunakan metode eulerian, pada posisi  $7.240659^\circ$  LS dan  $113.534143^\circ$  BT. Titik penelitian ditentukan berdasarkan pertimbangan lokasi perairannya yang dekat dengan tambak garam pesisir Desa Pademawu. Lokasi di titik tersebut memiliki profil pantai yang landai, serta termasuk pada zonasi laut dangkal. Titik koordinat di plot menjadi sebuah peta menggunakan *software* Quantum GIS (QGIS) versi 3.22.6. (Gambar 1). Metode euler diterapkan dengan peletakan instrumen pengukuran arus pada suatu titik di suatu kolom air secara temporal, sehingga data arus yang di dapatkan merupakan data berjenis *time series* atau dalam fungsi waktu (Permadi *et al.*, 2015). Adapun alur penelitian ini dapat dilihat pada Gambar 2.



**Gambar 1.** Lokasi Pengambilan Data Arus Laut



Gambar 2. Alur Penelitian



Gambar 3. Ilustrasi lapisan kedalaman pengukuran

Pengukuran instrumen diawali dengan peletakan ADCP di rangka ADCP pada dasar perairan, dengan sedimen berupa pasir. Rangka ADCP yang digunakan difungsikan supaya instrumen tetap stabil dan tidak terganggu oleh aktivitas arus dan gelombang. Perekaman ADCP dilakukan setiap 5 – 15 menit yang terbagi menjadi 10 lapisan kedalaman, dengan total kedalaman yaitu 10,4 m.

Profil kedalaman arus diurutkan dari lapisan 1 yang direkam pada lapisan paling bawah atau dasar laut, hingga lapisan 10 yang berada pada zona permukaan laut. Pembagian kedalaman didasarkan pada kedalaman lokasi kawasan penelitian yang mencapai 10,4 meter. Kedalaman dibagi kedalam tiap 1 m lapisan yaitu 1,4 m, 2,4 m, 3,4 m, 4,4 m, 5, 4m, 6,4 m, 7,4 m, 8,4 m, 9,4 m, dan 10,4 m, sehingga analisis yang dilakukan terhadap arus dapat lebih terperinci dalam mengamati dan mengetahui arah dan kecepatan arus dari tiap-tiap lapisan kedalaman (Bonauli, 2016). Ilustrasi lapisan kedalaman pengukuran dapat diamati berdasarkan Gambar (3). Data yang diperoleh dari instrumen ADCP memiliki format berbentuk binary (\*.ENR, \*.ENS, \*.ENX, \*.N1R, \*.STA, \*.LOG). Perubahan format menjadi ASCII-TXT (\*.TXT) diperlukan untuk membuka dan menampilkan data untuk diolah menggunakan bahasa program R.

Bahasa program R yang digunakan merupakan R versi 4.1.2. Bahasa program R diakses menggunakan *software* R-Studio. Pengolahan data melalui bahasa program R dilakukan dengan metode statistika deskriptif, dimana metode tersebut digunakan untuk menyederhanakan data agar mudah dipahami melalui *script* yang disusun bersama *packages* yang telah tersedia di R (Yanti 2021). Tahapan pengolahan data dilakukan dengan 3 (tiga) tahap, yaitu pre-pengolahan data, pengolahan data, dan post-pengolahan data. Tahapan tersebut dapat ditafsirkan sesuai dengan pemaparan Rule *et al.* (2018) dalam kajiannya yaitu seorang *data analyst* melibatkan beberapa siklus dalam menganalisis data diantaranya memperoleh, membersihkan, membuat profil, menganalisis, dan menafsirkan data. Siklus tersebut dirangkum dalam tahapan-tahapan yang dilakukan dalam penelitian ini.

## HASIL DAN PEMBAHASAN

Pada tahap awal dari pre-pengolahan data dilakukan *importing* data dan *Exploratory Data Analysis* (EDA). Data mentah ADCP merupakan

suatu *dataset* yang terdiri dari 451 *records* dan 23 *variable* yang dapat terekam oleh ADCP. Kolom pertama dalam data ADCP merupakan waktu dan tanggal perekaman (*date-time*), dimana pada kolom tersebut akan diketahui jarak tiap berapa menit ADCP melakukan perekaman. Pada kolom kedua dan ketiga merupakan tekanan (*pressure*), dan suhu (*temperature*). Kolom ke 4 hingga kolom ke 22 dengan loncatan 1 merupakan data kecepatan (*speed*) tiap lapisan kedalaman, sementara kolom ke 5 hingga kolom ke 23 dengan loncatan 2 merupakan arah (*direction*) tiap lapisan kedalaman.

EDA merupakan suatu seni dan ilmu dalam melakukan penyelidikan awal pada data melalui teknik statistik dan visualisasi yang dapat memunculkan aspek-aspek penting dalam data sehingga dapat digunakan untuk analisis lebih lanjut. Pembuatan hipotesis dan pengenalan pola pada data mentah ditekankan di tahap EDA (Putatunda *et al.*, 2019). Pengertian EDA juga dipaparkan Salgado (2016) sebagai tahap awal yang penting dalam melakukan suatu analisis data setelah tahap pengumpulan data, yang berguna untuk memberi arahan menuju hipotesis penelitian. Pengolahan data pada tahap EDA menggunakan R dilakukan dengan melakukan analisis dan pengenalan pola data, setelah dilakukan *importing* pada data mentah ADCP. Proses tersebut dijalankan dengan memunculkan 6 baris bagian teratas (Gambar 4) menggunakan fungsi `head()`, dan 6 baris terbawah (Gambar 5) dari data mentah ADCP menggunakan fungsi `tail()`, serta memeriksa struktur dari *dataset* (Gambar 6) menggunakan fungsi `str()`.

```

```{r}
ADCP <- read.csv("Data ADCP.csv", sep=";")
```
```{r}
head(ADCP)
tail(ADCP)
str(ADCP)
```

```

Langkah selanjutnya pada tahap pre-pengolahan data ialah dilakukannya deteksi *outlier* atau *noise*, dan data kosong atau *Non-Available Data* (NA). Keberadaan *outlier* umumnya disebabkan oleh beberapa hal, diantaranya kesalahan mekanis, perubahan perilaku sistem, aktivitas yang kurang ataupun tidak benar dalam pengambilan data, *human error* atau kesalahan manusia, kesalahan *entry* data, dan perubahan

lingkungan pengambilan data (Wang *et al.*, 2019). Laksono *et al.* (2015) mendefinisikan *outlier* sebagai penyimpangan ekstrem dari data, yang mengganggu struktur data yang sebenarnya. Fenomena palsu atau tidak asli akan tercemin pada rangkaian data (*dataset*) yang mengandung *outlier*, akibat nilai bias yang dihasilkan pada proses analisis dan pengolahan data tersebut. *Outlier* dapat mendeteksi suatu kesalahan penting dalam suatu *dataset* dengan menunjukkan suatu anomali pada *dataset* tersebut. Proses deteksi *outlier* pada R dapat diterapkan melalui suatu visualisasi grafis menggunakan beberapa visualisasi grafis yang disajikan dalam R, salah satunya yaitu *boxplots*. Pearson (2018), dalam bukunya menjelaskan suatu kutipan mengenai alasan penggunaan visualisasi

grafis dalam menganalisa suatu data, dimana manusia jauh lebih baik dalam melihat pola dalam grafis daripada dalam kumpulan angka yang besar.

Proses pembersihan data (*data cleaning*) merupakan salah satu bagian dari tahapan *quality control* (QC) data ADCP. Hal ini dikarenakan data sampah yang terdeteksi (*outlier*) terjadi karena pada saat proses peletakan ADCP di dasar laut, terkadang ADCP sudah mulai melakukan perekaman sebelum mencapai dasar laut. Perekaman tersebut terjadi secara tidak normal sehingga menyebabkan anomali dan *outlier* pada data hasil perekaman, terutama yang paling terlihat pada komponen tekanan yang dipengaruhi oleh kedalaman laut. Oleh karena itu dalam proses ini, data yang dilakukan deteksi merupakan komponen

| Description: df [6 x 23] |          |             |               |             |               |             |               |             |
|--------------------------|----------|-------------|---------------|-------------|---------------|-------------|---------------|-------------|
| DateTime                 | Pressure | Temperature | Speed.1.1.4m. | Dir.1.1.4m. | Speed.2.2.4m. | Dir.2.2.4m. | Speed.3.3.4m. | Dir.3.3.4m. |
| 1 13/11/2020 11:15       | 2.689    | 32.15       | 0.089         | 26.57       | 0.260         | 278.85      | 0.107         | 334.63      |
| 2 13/11/2020 11:35       | 2.566    | 32.00       | 0.089         | 45.00       | 0.357         | 304.05      | 0.164         | 292.19      |
| 3 13/11/2020 11:40       | 2.533    | 32.02       | 0.060         | 24.44       | 0.283         | 303.63      | 0.120         | 332.80      |
| 4 13/11/2020 11:45       | 2.504    | 32.06       | 0.098         | 35.47       | 0.261         | 293.03      | 0.141         | 255.67      |
| 5 13/11/2020 11:50       | 2.471    | 32.08       | 0.119         | 45.00       | 0.322         | 299.74      | 0.098         | 316.66      |
| 6 13/11/2020 11:55       | 2.435    | 32.11       | 0.128         | 35.43       | 0.240         | 285.49      | 0.022         | 296.57      |

Gambar 4. 6 (Enam) Baris Ter-atas Data ADCP

| Description: df [6 x 23] |          |             |               |             |               |             |               |             |
|--------------------------|----------|-------------|---------------|-------------|---------------|-------------|---------------|-------------|
| DateTime                 | Pressure | Temperature | Speed.1.1.4m. | Dir.1.1.4m. | Speed.2.2.4m. | Dir.2.2.4m. | Speed.3.3.4m. | Dir.3.3.4m. |
| 446 15/11/2020 12:50     | 0.118    | 33.43       | 3.277         | 113.37      | 0.845         | 230.09      | 0.564         | 279.60      |
| 447 15/11/2020 12:55     | 0.117    | 35.06       | 2.066         | 101.42      | 0.977         | 39.64       | 0.234         | 192.85      |
| 448 15/11/2020 13:00     | 0.116    | 34.10       | 1.104         | 346.64      | 1.039         | 284.61      | 1.246         | 56.59       |
| 449 15/11/2020 13:05     | 0.116    | 33.94       | 3.026         | 276.78      | 0.821         | 297.25      | 0.722         | 152.09      |
| 450 15/11/2020 13:10     | 0.116    | 33.85       | 3.102         | 294.57      | 0.344         | 169.11      | 0.568         | 145.69      |
| 451 15/11/2020 13:15     | 0.115    | 33.83       | 2.612         | 30.31       | 0.404         | 2.27        | 0.552         | 20.57       |

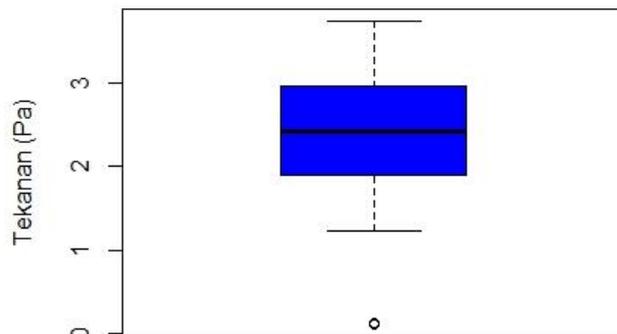
Gambar 5. 6 (Enam) Baris Ter-bawah Data ADCP

```

'data.frame': 451 obs. of 23 variables:
 $ DateTime : chr "13/11/2020 11:15" "13/11/2020 11:35" "13/11/2020 11:40"
 "13/11/2020 11:45" ...
 $ Pressure : num 2.69 2.57 2.53 2.5 2.47 ...
 $ Temperature : num 32.1 32 32 32.1 32.1 ...
 $ Speed.1.1.4m. : num 0.089 0.089 0.06 0.098 0.119 0.128 0.07 0.066 0.047 0.113 ...
 $ Dir.1.1.4m. : num 26.6 45.24 4.35 5.45 ...
 $ Speed.2.2.4m. : num 0.26 0.357 0.283 0.261 0.322 0.24 0.296 0.253 0.281 0.24 ...
 $ Dir.2.2.4m. : num 279 304 304 293 300 ...
 $ Speed.3.3.4m. : num 0.107 0.164 0.12 0.141 0.098 0.022 0.15 0.144 0.075 0.16 ...
 $ Dir.3.3.4m. : num 335 292 333 256 317 ...
 $ Speed.4.4.4m. : num 0.16 0.177 0.172 0.171 0.233 0.145 0.185 0.196 0.228 0.323 ...
 $ Dir.4.4.4m. : num 300 307 332 287 310 ...
 $ Speed.5.5.4m. : num 0.152 0.187 0.093 0.123 0.097 0.05 0.057 0.066 0.071 0.057 ...
 $ Dir.5.5.4m. : num 275 298 335 265 312 ...
 $ Speed.6.6.4m. : num 0.159 0.237 0.306 0.271 0.243 0.166 0.244 0.219 0.215 0.056
 ...
 $ Dir.6.6.4m. : num 134 132 123 117 115 ...
 $ Speed.7.7.4m. : num 0.057 0.177 0.065 0.204 0.154 0.197 0.1 0.152 0.236 0.247 ...
 $ Dir.7.7.4m. : num 176 273 302 249 280 ...
 $ Speed.8.8.4m. : num 0.208 0.269 0.123 0.121 0.132 0.104 0.062 0.11 0.107 0.096 ...
 $ Dir.8.8.4m. : num 278 285 302 252 257 ...
 $ Speed.9.9.4m. : num 0.077 0.031 0.126 0.094 0.066 0.076 0.027 0.031 0.183 0.135
 ...
 $ Dir.9.9.4m. : num 309.7 309.8 33.7 215 248.8 ...
 $ Speed.10.10.4m. : num 0.203 0.183 0.038 0.257 0.198 0.18 0.157 0.119 0.249 0.131 ...
 $ Dir.10.10.4m. : num 235 264 317 258 293 ...
    
```

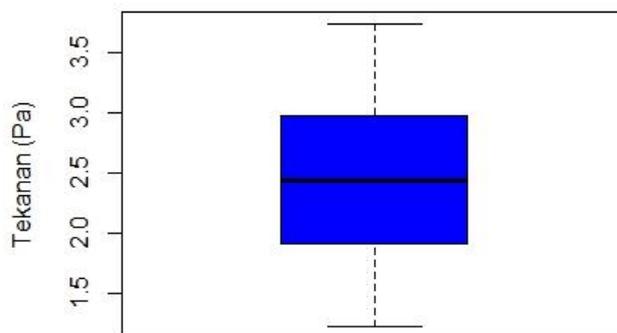
Gambar 6. Struktur *Dataset* ADCP

**Boxplot Data Tekanan ADCP**



**Gambar 7.** Boxplot Data Tekanan ADCP

**Boxplot Data Tekanan ADCP**



**Gambar 8.** Boxplot Data Tekanan ADCP

tekanan. Proses pembersihan data dilakukan dengan menghapus seluruh baris yang mengandung *outlier* pada data tekanan, sehingga melibatkan data waktu dan tanggal perekaman, suhu, kecepatan dan arah arus yang terdapat di baris yang sama pada tabel data. Visualisasi *boxplots* untuk menginterpretasikan ada tidaknya *outlier* dan data sampah dapat dilihat berdasarkan Gambar (7). Data mentah ADCP berdasarkan Gambar (7). Chu *et al.* (2016) menjelaskan dalam kajiannya bahwa data kotor dan yang belum melalui pembersihan dapat menimbulkan kesimpulan yang salah dan analisis yang tidak dapat diandalkan, terkait hal tersebut pembersihan data menjadi kunci utama dalam penelitian basis data. *Dataset* hasil pembersihan didapatkan 442 *records* dan 23 *variables*. Hasil pembersihan

*outlier* dapat diamati berdasarkan visualisasi ulang pada Gambar (8).

**Pengolahan Data**

Pengolahan data menggunakan R pada tahap pengolahan data dilakukan ketika data ADCP sudah melalui *cleaning data*. Pada tahapan ini akan dihasilkan suatu *dataset* yang memisahkan arah dan kecepatan arus komponen Timur-Barat (U) dan arus komponen Utara-Selatan (V), serta arus pasut/ arus harmonik dan arus non-pasut/ arus non-harmonik/ arus residu. Simatupang *et al.* (2016) juga melakukan pemisahan arus dalam penelitiannya, di perairan muara Sungai Banyuasin, Provinsi Sumatera Utara. Pemisahan arus dalam penelitian tersebut dilakukan serupa dengan penelitian ini yaitu setelah tahap pre-pengolahan data. Pemisahan dan penguraian arus

menjadi sejumlah komponen perlu dilakukan untuk membantu pengklasifikasian sistem sirkulasi arus (Theoyana *et al.*, 2015). Proses ini dikalkulasi menggunakan beberapa rumus yang telah dipaparkan oleh Pratama *et al.* (2016) dalam kajiannya, yang disusun dengan pengkodean menggunakan bahasa program R dengan melalui tahapan, diantaranya:

- Mengubah arah/*direction* ke radian (Q)
- Mencari kecepatan total ( $V_{tot}$ )
- Mencari kecepatan arus komponen timur-barat ( $v_x$ ) dan kecepatan arus komponen utara-selatan ( $v_y$ )
- Mencari rata-rata kecepatan arus komponen timur-barat ( $vr_x$ ) dan rata-rata kecepatan arus komponen utara-selatan ( $vry$ )
- Mencari  $\sum v_x - vr_x$  (A)
- Mencari  $\sum v_y - vry$  (B)

Proses pengolahan data dilanjutkan dengan tahapan mencari nilai arah dan kecepatan arus pasut dan arus non-pasut. Tahap ini diproses berdasarkan rumus-rumus yang dipaparkan oleh Simatupang *et al.* (2016) dalam kajiannya, serta komponen pengolahan arus yang telah dihasilkan sebelumnya.

```

```{r}
#Mencari kecepatan (v) arus pasut
vpasut <- sqrt(A^2+B^2)
# Mencari kecepatan (v) arus non-pasut
vnonpasut <- vtot-vpasut
#Mencari arah (u) arus pasut
Upasut <- atan(B/A)
# Mencari arah (u) arus non-pasut
Unonpasut <- atan(vrx/vry)
```

```

Berdasarkan perhitungan menggunakan rumus yang telah diolah menggunakan bahasa R, diketahui bahwa Perairan Pademawu didominasi oleh arus pasut V (Utara – Selatan) dengan kecepatan rata-rata sebesar 0,11 – 0,181 m/s. Rata-rata kecepatan komponen arus pasut U (Timur-Barat) dan V (Utara – Selatan) dapat diamati berdasarkan Tabel (1).

### Post-Pengolahan data

Karakteristik arus yang dapat diinterpretasikan melalui stick plot merupakan arah dan kecepatan dari arus pada lapisan 1 yang berada pada dasar perairan, hingga lapisan 10 yang berada pada permukaan perairan. Visualisasi pada stick plot ditampilkan pada tiap lapisan kedalaman secara *time series* berdasarkan Gambar (8). Proses

visualisasi dilakukan dengan menggunakan pengkodean di R dengan *package* “oce” melibatkan arus komponen Timur-Barat (U) dan Selatan-Utara (V).

```

```{r}
library(oce)
```{r}
PS1 <- plotSticks(PS$DateTime,
0,PS$Speed.1.1.4m.x , PS$Speed.1.1.4m.y ,
yscale=10, xlab= "waktu", ylab=
"Kecepatan (m/s)",ylim= c(-0.5,0.5),
xaxt="n", length= 0.5)
title(main = "Lapisan 1", cex.main= 1,
font.main= 3)
axis.POSIXct(1, at=PSS$DateTime,
labels=format(PSS$DateTime, "%d/%m/%Y
%H:%M"))

```

Arah arus berdasarkan stick plot dapat dianalisis berdasarkan arah mata angin. Interpretatif *stick plot* arah utara ditunjukkan dengan orientasi garis *stick plot* ke atas ( $0^\circ$ ), timur  $90^\circ$ , selatan  $180^\circ$ , barat  $270^\circ$  dengan arah putaran searah jarum jam (Naulita, 2016). Berdasarkan visualisasi *stick plot* tersebut diketahui bahwa arus V (Selatan-Utara) mendominasi perairan penelitian pada 10 lapisan kedalaman di bulan November.

Visualisasi juga ditampilkan dalam bentuk grafik *time series* arus komponen U (Timur-Barat) dan V (Selatan-Utara). Grafik U dan V divisualisasikan pada 10 lapisan kedalaman untuk dapat menghasilkan informasi mengenai karakteristik arus di area penelitian secara vertikal. Hasil grafik dapat amati berdasarkan Gambar (9). Pengkodean grafik U dan V dilakukan dengan menggunakan R *base*.

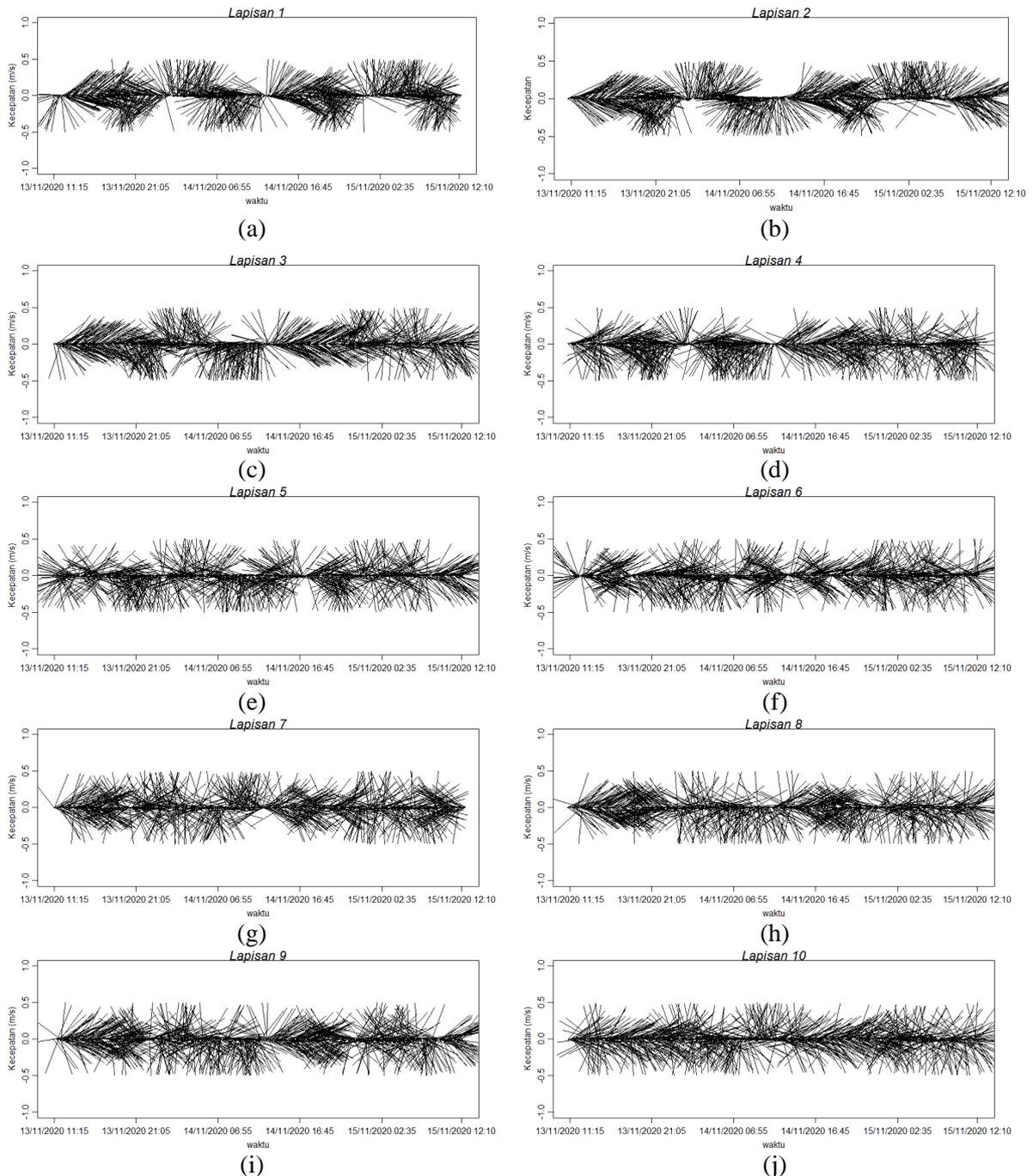
```

```{r}
G1 <- plot(PSS$DateTime,
PSS$Speed.1.1.4m.x,
type="l", col=2, ylim=c(-0.5,0.5),
xlab="Waktu", ylab="Kecepatan (m/s)",
xaxt="n")
lines(PSS$DateTime, PSS$Speed.1.1.4m.y,
type="l", col=3)
title(main = "Lapisan 1", cex.main= 1,
font.main= 3)
legend("topright", c("Arus U", "Arus V"), lty =
1, col = 2:4, cex = 1)
axis.POSIXct(1, at=PSS$DateTime,
labels=format(PSS$DateTime, "%d/%m/%Y
%H:%M"))
```

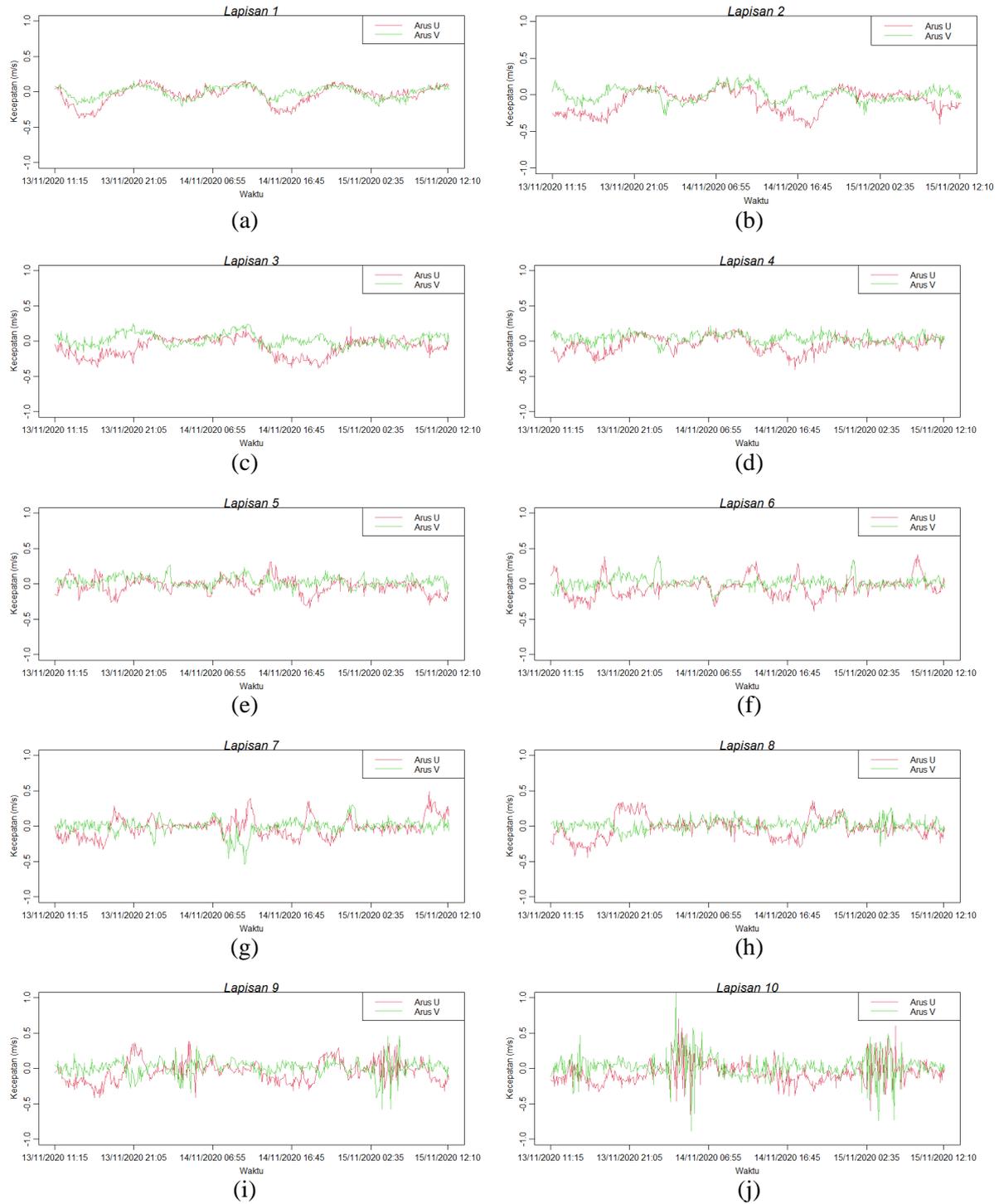
```

**Tabel 1.** Rata-rata kecepatan komponen arus.

| Jenis Arus | Rata-Rata Kecepatan Arus V (m/s) | Rata-Rata Kecepatan Arus U (m/s) |
|------------|----------------------------------|----------------------------------|
| Pasut      | 0.11 - 0.181                     | (-0.031) - (-0.021)              |
| Non Pasut  | 0.004 - 0.016                    | (-0.682) - (-1.24)               |



**Gambar 9.** Visualisasi *Time Series Stick Plot* Menggunakan Bahasa Program R A) Lapisan 1 (10,4 m), B) Lapisan 2 (9,4 m), C) Lapisan 3 (8,4 m), D) Lapisan 4 (7,4m), E) Lapisan 5 (6,4m), F) Lapisan 6 (5,4 m), G) Lapisan 7 (4,4 m), H) Lapisan 8 (3,4 m), I) Lapisan 9 (2,4 m), J) Lapisan 10 (1,4 m).



**Gambar 10.** Grafik garis *Time Series* U dan V Menggunakan Bahasa Program R A) Lapisan 1(10,4 m), B) Lapisan 2 (9,4 m), C) Lapisan 3 (8,4 m), D) Lapisan 4 (7,4m), E) Lapisan 5 (6,4m), F) Lapisan 6 (5,4 m), G) Lapisan 7 (4,4 m), H) Lapisan 8 (3,4 m), I) Lapisan 9 (2,4 m), J) Lapisan 10 (1,4 m).

Berdasarkan hasil analisa data pengukuran arus ADCP diketahui bahwa kecepatan arus rata-rata berkisar 0.114-0.197 m/s. Kecepatan arus

maksimum berada di lapisan 2 yaitu sebesar 0,456 m/s, sedangkan kecepatan arus minimum berada di arus lapisan 3 yaitu sebesar 0,004 m/s.

## KESIMPULAN

Tahapan pengolahan data mentah ADCP menggunakan bahasa program R diantaranya yaitu pre-pengolahan, pengolahan, dan post-pengolahan data. Pada tahap pre-pengolahan data, data mentah ADCP melewati proses EDA dan *cleaning data*, sehingga dihasilkan 224 *records* dan 23 *variable dataset* bersih. Pengolahan data dilakukan setelahnya dengan melibatkan algoritma yang dibutuhkan untuk pemisahan komponen arus. Data ADCP dapat divisualisasikan ketika melalui tahap post-pengolahan data. Pada tahap tersebut, dihasilkan 442 *records* dan 21 *variable dataset* sebagai bahan visualisasi *stick plot* dan grafik garis (*line chart*) U–V. Visualisasi menggunakan bahasa program R dapat dilakukan dengan sederhana menggunakan *package* yang tersedia, dengan menerapkan *syntax* yang sama pada tiap *layer* kedalaman. Output dari visualisasi tersebut berupa interpretasi karakteristik arus laut di kawasan Perairan Pademawu. Pengolahan data arus pada penelitian ini belum memaksimalkan penggunaan paket-paket / *packages* yang tersedia di bahasa program R. Oleh karena itu penelitian lebih lanjut diperlukan terkait hal tersebut.

## DAFTAR PUSTAKA

- Anisa, M. N., Purwanto, P., & Indra, I.B. 2017 Studi Pola Arus Laut Di Perairan Tapaktuan, Aceh Selatan, *Journal of Oceanography*, 6(1):183–192.
- Asatidz, S., Alfi, S., Aris, I., Haryoso, S., & Purwanto. 2021. Pemodelan Sebaran Sedimen Dasar di Perairan Pelabuhan Branta Pamekasan, *Indonesian Journal of Oceanography*, 3(1):64-75.
- Budiharto, W. & Rachmawati, R.N. 2013 Pengantar Praktis Pemrograman R, pp. 3–13.
- Cholissodin, I., & Riyandani, E. 2016. Analisis Big Data. Fakultas Ilmu Komputer (Fikom), Universitas Brawijaya (UB), Malang.
- Bonauli, M., Muhammad, H., & Widodo, S.P. 2016. Analisis Karakteristik Arus Harmonikakibat Pasang Surut Di Perairan Teluk Awur Kabupaten Jepara, *Jurnal Oseanografi*, 5(1):1-10.
- Chu, X., Ilyas, I.F., Krishnan, S., & Wang. 2016 Data cleaning: Overview and emerging challenges, *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 26-June-20, pp. 2201–2206. doi: 10.1145/2882903.2912574.
- Dewi, I.P., Nursalam, N. & Widyanata, D. 2021. Pengaruh Dinamika Oseanografi Terhadap Ekosistem Mangrove Di Desa Pagatan Besar, *Jurnal Geoelebes*, 5(1):35–45. doi: 10.20956/geoelebes.v5i1.11975.
- Gatto, L., Lisa., M. B., Thomas, N., & Sebastian, G. 2015. Visualization of proteomics data using R and Bioconductor. *Proteomics*, 15(8): 1375–1389. doi: 10.1002/pmic.2014003
- Hu, Y. & Yang, J. 2014. Role of ocean heat transport in climates of tidally locked exoplanets around M dwarf stars, *Proceedings of the National Academy of Sciences of the United States of America*, 111(2):629–634. doi: 10.1073/pnas.1315215111.
- Indrayanti, E., Wijayanti, D.P. & Siagian, H.S.R. 2020. Pasang Surut, Arus dan Gelombang Berdasarkan Data Pengukuran Acoustic Doppler Current Profiler di Perairan Pulau Cilik, Karimunjawa, *Buletin Oseanografi Marina*, 9(1):37–44. doi: 10.14710/buloma.v9i1.29065.
- Kaya, E., Agca, M., Adiguzel, F., & Cetin, M. 2019. Spatial data analysis with R programming for environment, *Human and Ecological Risk Assessment: An International Journal*, 25(6):1521–1530.
- Kumagai, N.H., García Molinos, J., Yamano, H., Takao, S., Fujii, M., & Yamanaka, Y. 2018 ‘Ocean currents and herbivory drive macroalgae-to-coral community shift under climate warming. *Proceedings of the National Academy of Sciences of the United States of America*, 115(36):8990–8995. doi: 10.1073/pnas.1716826115.
- Laksono, M.A.T., Yudha, P., & Astri, N. 2015 Analisis Sistem Deteksi Anomali Trafik Menggunakan Algoritma Clustering Cure (clustering Using Representatives) dengan Outlier Removal Clustering dalam Menangani Outlier. *eProceedings of Engineering*, 2(2): 3800-3807.
- Lemenkova, P. 2018. Statistical Analysis Of The Mariana Trench Geomorphology Using R Programming Language, *Geodesy and Cartography*, 45(2):57-84
- Liu, Y., Meng, Q., Chao, L., & Zhongwen, G. 2017 Big data challenges in ocean observation: a survey, *Personal and Ubiquitous Computing*, 21(1):55–65. doi: 10.1007/s00779-016-0980-2.
- Naulita, Y. 2016. Proses pencampuran turbulen di kanal Labani, Selat Makassar. *Jurnal Ilmu dan Teknologi Kelautan Tropis*, 8(1):345-355.

- Pearson, R. K. 2018 *Exploratory Data Analysis Using R*, CRC Press Taylor & Francis Group Chapman and Hall Book.
- Permadi, L.C., Indrayanti, E. & Rochaddi, B. 2015. Studi Arus Pada Perairan Laut Di Sekitar Pltu Sumuradem, *Jurnal Oseanografi*, 4(2):516–523.
- Pratama, D.R., Yusuf, M. & Helmi, M. 2016 Kajian Kondisi Dan Sebaran Kualitas Air Di Perairan Selatan Kabupaten Sampang, Provinsi Jawa Timur. *Jurnal Oseanografi*, 5:479–488.
- Putatunda, S., Rama, K., Ubrangala, D., dan Kondapalli, R. 2019. SmartEDA: An R Package for Automated Exploratory Data Analysis. *Journal of Open Source Software*, 4(41): 1-18. doi:10.21105/joss.01509.
- Rasheva-Yordanova, K., Veselin, C., Iva, K., Evtim, I., Pepa, P., & Boriana, N. 2018. Forming of Data Science Competence for Bridging the Digital Divide. *International Conference The Future of Education*, June, p.ENT3591
- Rule, A., Tabard, A. and Hollan, J.D., 2018, April. Exploration and explanation in computational notebooks. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12. doi: 10.1145/3173574.3173606.
- Salgado, C.M. 2016. Setting the Stage: Rationale Behind and Challenges to Health Data Analysis, *Secondary Analysis of Electronic Health Records*. Cham (CH): Springer
- Sari, K.A.D.P., Suryandari, N.N.A. dan Putra, G.B.B. 2021. Pengaruh Pemanfaatan Teknologi, Partisipasi Pemakai, Kemampuan Teknik Pemakai, Pengalaman Kerja dan Jabatan Terhadap Efektivitas Sistem Informasi Akuntansi. *Kharisma*, 3(1):1–11.
- Simatupang, C.M. & Agussalim, A. 2016. Analisis Data Arus Di Perairan Muara Sungai Banyuasin Provinsi Sumatera Selatan Analysis Of Flow Data On Estuarine Banyuasin River In South Sumatera, *Maspari Journal*, 8(1):15–24.
- Siregar, L.Y. & Nasution, M.I.P. 2020. Perkembangan Teknologi Informasi Terhadap Peningkatan Bisnis Online. *Hirarki : Jurnal Ilmiah Manajemen dan Bisnis*, 02(01):71–75.
- Tanto, A.T., Wisna, U.J., Kusumah, G., Pranowo, W.S., Husrin, S., Ilham, I., & Putra, A. 2017. Karakteristik Arus Laut Perairan Teluk Benoa – Bali. *Jurnal Ilmiah Geomatika*, 23(1):37-48. doi: 10.24895/jig.2017.23-1.631.
- Theoyana, T.A., Widodo, S.P., Anastrasia, R. T. D. K., & Purwanto. 2015. Karakteristik arus pasang surut di selat badung, bali. *Jurnal Segara*, 11(2):115–123.
- Thimbleby, H. 2013. Technology and the Future of Healthcare. *Journal of Public Health Research*, 2(3): 160-167. doi: 10.4081/jphr.2013.e28.
- Wang, H., Bah, M.J. & Hammad, M. 2019. Progress in Outlier Detection Techniques: A Survey. *IEEE Access*, 7, pp. 107964–108000. doi: 10.1109/ACCESS.2019.2932769.
- Wattimena, M.C., Atmadipoera, A.S. & Purba, M., Koch-Larrouy, A. 2014. Variabilitas Intra-Musiman Arus Dekat-Dasar Di Laut Halmahera. *Jurnal Ilmu dan Teknologi Kelautan Tropis*, 6(2):267–281.
- Wisha, U.J., Husrin, S. & Prihantono, J. 2015. Hydrodynamics Banten Bay During Transitional Seasons (August-September) (Hidrodinamika Perairan Teluk Banten Pada Musim Peralihan (Agustus–September)). *Ilmu Kelautan: Indonesian Journal of Marine Sciences*, 20(2):101-112. doi: 10.14710/ik.ijms.20.2.101-112.
- Yanti, H.A. 2021. Pengolahan Data Sederhana Menggunakan R Studio. *Jurnal Sienna*, 2(1):1-9
- Yudistira, I.G.A.A. 2013. Desain Animasi Statistika Berbasis Bahasa Pemrograman R. *E-Journal Widya Eksakta*, 1(1):1-6.
- Yuniarti, R. 2022. Kesalahan Mahasiswa Program Studi Administrasi Publik Dalam Menyelesaikan Soal Statistika Deskriptif Dan Statistika Inferensial, *Jurnal Sains Matematika dan Statistika*, 8(1):46-58. doi: 10.24014/jsms.v8i1.13312.