

BAB I PENDAHULUAN

1.1 Latar Belakang

Indonesia, sebagai negara kepulauan, memiliki sebagian besar wilayah berupa lautan. Oleh sebab itu, berbagai kegiatan di laut seperti pelayaran dan penangkapan ikan menjadi aspek vital bagi kehidupan masyarakat Indonesia. Semua aktivitas kelautan sangat bergantung dan rentan terhadap perubahan yang terjadi di laut. Perairan Indonesia memiliki potensi produksi ikan laut berkelanjutan yang cukup besar, yakni sekitar 6,51 juta ton per tahun, atau sekitar 8,2% dari total potensi produksi ikan laut dunia. Ikan pelagis umumnya hidup bergerombol, baik dengan spesies yang sama maupun berbeda, meskipun mereka cenderung membentuk kelompok berdasarkan ukuran tubuh. Dari sisi klimatologi, wilayah Indonesia dipengaruhi oleh angin musim barat dan timur, di mana dinamika angin ini berdampak langsung terhadap perubahan yang terjadi di perairan. Kondisi monsun di perairan Indonesia mencerminkan adanya interaksi reguler antara laut dan atmosfer lokal (Suwardi dkk., 2023).

Kegiatan penangkapan ikan merupakan upaya memperoleh hasil tangkapan guna memenuhi kebutuhan konsumsi masyarakat dengan memanfaatkan berbagai jenis alat tangkap. Permintaan terhadap hasil tangkapan tersebut mendorong terbentuknya siklus ekonomi yang melibatkan potensi keuntungan maupun kerugian. Oleh sebab itu, pelaku usaha penangkapan ikan berusaha meningkatkan volume produksi untuk mengoptimalkan perolehan keuntungan. Namun demikian, peningkatan produksi tersebut sangat bergantung pada kondisi lokasi penangkapan, yang dipengaruhi oleh beragam faktor yang saling berinteraksi (Suwardi dkk., 2023).

Salah satu sumber daya ikan yang memiliki potensi besar di Indonesia adalah kelompok ikan pelagis, di antaranya cakalang, tuna, dan layang. Indonesia memainkan peran yang signifikan dalam produksi global cakalang, tuna, dan tongkol. Pada tahun 2011, produksi global untuk ketiga jenis ikan tersebut mencapai 6,8 juta ton dan meningkat menjadi lebih dari 7 juta ton pada tahun 2012, dengan rata-rata produksi selama periode 2005–2012 sebesar 1.033.211 ton. Indonesia berkontribusi lebih dari 16% terhadap total produksi dunia untuk cakalang, tuna, dan tongkol. Pada tahun 2013, volume ekspor ketiga komoditas sekitar 209.410 ton dengan nilai ekspor sebesar USD 764,8 juta. Indonesia juga tercatat sebagai negara dengan kontribusi terbesar di antara 32 negara anggota *Indian Ocean Tuna Commission* dengan rata-rata produksi sebesar 356.862 ton per tahun selama periode



Berdasarkan uraian tersebut, peneliti terdorong untuk melakukan peramalan hasil tangkapan ikan dengan mempertimbangkan faktor-faktor yang telah disebutkan sebelumnya. Penelitian ini difokuskan pada data dengan pola triwulanan, sehingga metode yang digunakan meliputi *Random Forest*, *Linear Regression*, dan *k-fold Cross Validation*. Penerapan metode-metode ini diharapkan dapat meningkatkan tingkat akurasi dalam proses prediksi. Adapun parameter yang digunakan dalam penelitian ini meliputi ENSO, IOD, MJO, suhu permukaan laut, salinitas, dan kecepatan arus.

1.2 Tujuan dan Manfaat

1.2.1 Tujuan

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis pengaruh variabel iklim terhadap hasil tangkapan ikan cakalang (*Katsuwonus pelamis*) di perairan Indonesia.
2. Mengembangkan model prediksi hasil tangkapan menggunakan metode Linear Regression dan Random Forest.
3. Mengevaluasi akurasi model dengan teknik k-Fold Cross Validation.

1.2.2 Manfaat

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Bagi Sektor Perikanan Tangkap
Penelitian ini dapat memberikan acuan prediksi hasil tangkapan ikan cakalang berdasarkan dinamika iklim laut, yang berguna dalam menentukan strategi penangkapan yang lebih tepat sasaran. Model yang dikembangkan diharapkan mampu meningkatkan efisiensi operasional dan meminimalkan ketidakpastian hasil tangkapan.
2. Bagi Lembaga Meteorologi
Penelitian ini dapat mendukung lembaga meteorologi dalam memahami dan memetakan dampak variabilitas iklim laut terhadap sektor perikanan. Model prediksi yang dibangun berbasis data klimatologis dapat digunakan sebagai salah satu komponen dalam pengembangan sistem informasi iklim kelautan, yang mendukung pengambilan keputusan lintas sektor, termasuk mitigasi risiko pada aktivitas maritim dan perikanan.
3. Bagi Peneliti dan Akademisi
Penelitian ini memberikan kontribusi dalam pengembangan ilmu pengetahuan di bidang klimatologi perairan dan prediksi hasil perikanan.



Metode Random Forest, Linear Regression, dan validasi k-Fold Cross Validation dalam konteks ini diharapkan dapat menghasilkan metodologis untuk penelitian-penelitian sejenis di masa

1.3 Landasan Teori

1.3.1 Ikan Cakalang

Ikan cakalang atau dalam Bahasa Inggris disebut skipjack tuna adalah salah satu ikan yang umum dan sering ditangkap oleh para nelayan serta termasuk ke dalam keluarga ikan tuna namun cenderung lebih kecil. Ikan Cakalang (*Katsuwonus pelamis*) merupakan salah satu jenis ikan pelagis besar yang banyak tertangkap di perairan Indonesia. Ikan cakalang sering membentuk schooling di sekitar permukaan dan schooling ini dapat diketahui dengan memperhatikan tanda-tanda alam seperti burung-burung yang terbang rendah, benda-benda terapung, hiu dan paus, serta sering menunjukkan tingkah laku yang unik dengan cara meloncat ke udara, memburu mangsa, membentuk buih, dan lain-lain (Fathaero, 2020).



Gambar 1. Ikan Cakalang

Sumber: fishbase, 2025

Filum : *Chordata*

Kelas : *Pisces*

Ordo : *Perciformes*

Famili : *Scombridae*

Subfamili : *Thunninae*

Genus : *Katsuwonus*



us pelamis

memiliki tubuh yang padat, penampang bulat, *lateral line* yang terletak di bawah sirip punggung kedua, sirip dada pendek tiga. Pada perairan Indonesia terdapat hubungan yang nyata cakalang dengan ikan pelagis kecil serta plankton. Semakin

banyaknya ikan kecil dan plankton, maka cakalang akan berkumpul untuk mencari makan. Ikan cakalang biasanya akan membentuk gerombolan (schooling) pada saat ikan tersebut mencari makanan. Ikan cakalang sering membentuk schooling di sekitar permukaan dan schooling ini dapat diketahui dengan memperhatikan tanda-tanda alam seperti burung-burung yang terbang rendah, benda-benda terapung, hiu dan paus, serta sering menunjukkan tingkah laku yang unik dengan cara meloncat ke udara, memburu mangsa, membentuk buih, dan lain-lain (Ismunandar dkk, 2018).

1.3.2 ENSO (El Niño-Southern Oscillation)

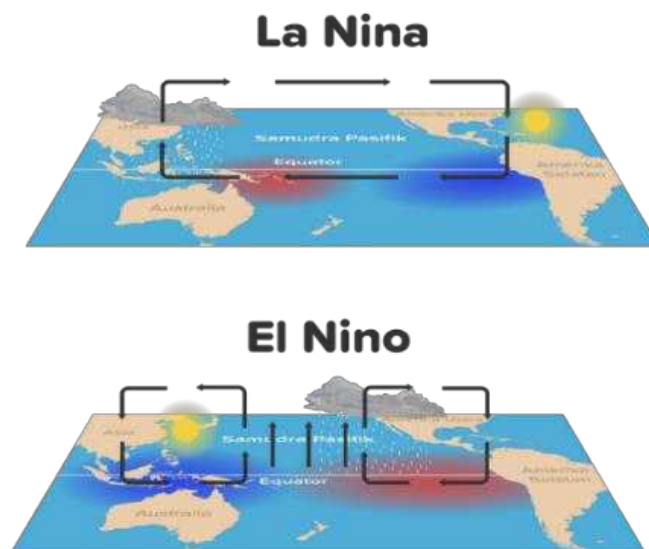
ENSO merupakan fenomena laut-atmosfer yang dipengaruhi oleh interaksi kompleks dari variabel iklim, seperti awan, angin, suhu permukaan laut, dan arus di Samudra Pasifik yang memberikan dampak yang berbedadi berbagai belahan dunia. ENSO dibagi menjadi: kondisi El Nino, La Nina dan kondisi Normal. ENSO merupakan sebuah interaksi laut atmosfer yang berpusat di wilayah ekuator Samudra Pasifik (Aldrian, 2008) yang menyebabkan anomali iklim global (Trenberth & Caron, 2000). Gejala ENSO memberikan pengaruh terhadap kondisi laut di Indonesia yaitu menjadi lebih dingin pada tahun El Nino dan lebih hangat pada tahun La Nina (Aldrian, 2008).

Anomali positif suhu muka laut (El Niño) dapat menyebabkan kemarau ekstrem sedangkan anomali negatif suhu muka laut (La Niña) dapat menyebabkan musim hujan berkepanjangan di sebagian besar wilayah Indonesia. El Niñomerupakan fase positif dari ENSO. El Niño terjadi ketika *sea surface temperature* (SST) di Pasifik tropis tengah dan timur lebih hangat dari rata-rata. Sedangkan La Niña merupakan fase negatif dari ENSO. La Niña terjadi ketika SST di Pasifik tropis bagian tengah dan timur lebih dingin dari rata-rata. Dampaknya di sebagian besar wilayah Indonesia, musim kemarau menjadi sangat kering serta permulaan musim hujan yang terlambat ketika El Niño berlangsung. Sedangkan ketika La Niña, musim penghujan akan tiba lebih awal dari biasanya. ENSO merupakan fenomena gabungan interaksi lautan atmosfer yang menyebabkan variasi iklim tahunan di dunia (Ropelewski dan Halpert, 1987; Trenberth dan Caron, 2000; Trenberth dan Stephaniak, 2001; Mc Bride dkk., 2003; Aldrian, 2008). Komponen lautan dari ENSO adalah El Nino dan La Nina, sedangkan komponen atmosfernya adalah Southern Oscillation (Trenberth, 1997). Pusat aktivitas ENSO berada di Samudra Pasifik dengan garis ekuator (Trenberth, 1997; Trenberth dan



igambarkan untuk menggambarkan keadaan tahunan arus ng menyusuri pantai selatan Peru dan Ekuador yang nya tangkapan ikan (Trenberth, 1997). Turunnya tangkapan abkan oleh nutrisi yang biasanya dimunculkan ke permukaan

oleh umbalan (*upwelling*) melemah. Perkembangan selanjutnya El Nino merupakan keadaan peningkatan suhu permukaan lautan (*sea surface temperature*) dari suhu normalnya di Pasifik Ekuator timur. La Nina adalah kejadian berkebalikan dengan El Nino yakni penurunan suhu permukaan lautan di kawasan ekuator Samudera Pasifik dari suhu normalnya (Trenberth, 1997; Prabowo, 2002). Ketika terjadi El Nino maupun La Nina, keduanya berasosiasi dengan Southern Oscillation, sehingga fenomena ini lebih dikenal sebagai ENSO (Trenberth, 1997; Aldrian, 2008; Prabowo, 2002).



Gambar 2.Perbandingan La Nina dan El Nino

Sumber: <https://roboguru.ruangguru.com/question/jelaskan-kaitan-antara-fenomena-el-nino-dan-la-nina-dengan-kenaikan-harga> QU-RFDI24D3

1.3.3 MJO (*Madden Julian Oscillation*)

Salah satu fenomena global yang turut memengaruhi pola cuaca dan iklim di wilayah Indonesia adalah *Madden Julian Oscillation* (MJO). Menurut penelitian oleh Madden & Julian (1971), MJO merupakan bentuk osilasi utama yang terjadi dalam dinamika iklim kawasan tropis. Istilah osilasi merujuk pada fluktuasi berkala suatu parameter terhadap waktu. Pengaruh MJO paling dominan di lintang rendah, terutama di sekitar garis khatulistiwa.



ama kali muncul di wilayah Samudra Hindia dan bergerak arah timur, dalam rentang lintang antara 100° LU hingga 100° siklus berkisar antara 30 hingga 60 hari dan ditandai oleh ulusi atmosfer berskala besar yang disertai dengan aktivitas

konveksi yang intens. Pergerakannya dimulai dari Samudra Hindia di sebelah barat Indonesia menuju Samudra Pasifik di sebelah timur, dengan kecepatan sekitar 5 meter per detik.

MJO memiliki peran penting dalam menjelaskan variasi iklim di wilayah tropis dan erat kaitannya dengan pembentukan area perairan hangat (*warm pool*) di timur Samudra Hindia serta barat Samudra Pasifik. Selama bergerak ke arah timur, MJO membawa angin baratan (*westerly wind*) sepanjang zona ekuator, yang disertai dengan pertumbuhan awan kumulonimbus konvektif yang padat. Awan-awan ini menyebabkan curah hujan lebat di sepanjang jalur pergerakan MJO. Di wilayah Samudra Hindia, fenomena ini menjalar dengan kecepatan sekitar 100 kilometer per hari, dan meningkat hingga 500 kilometer per hari saat melintasi wilayah Indonesia (Evana dkk., 2008).

1.3.4 IOD (*Indian Ocean Dipole*)

Pergeseran musim di Indonesia dapat dipengaruhi oleh salah satu fenomena oseanografi penting, yaitu *Indian Ocean Dipole (IOD)*. IOD terjadi ketika terdapat ketimpangan anomali suhu permukaan laut antara bagian barat dan timur Samudra Hindia (Aldrian, E., 2008; Saji dkk., 1999). Perbedaan ini menyebabkan perubahan suhu di permukaan laut yang selanjutnya dapat berdampak pada naik atau turunnya intensitas curah hujan di wilayah terdampak.

Secara geografis, Indonesia berada di antara dua benua Asia dan Australia serta diapit oleh dua samudra besar, yakni Samudra Hindia dan Samudra Pasifik. Letak ini menjadikan wilayah perairan Indonesia sangat rentan terhadap dinamika oseanografi kedua samudra, terutama Samudra Hindia yang kerap menjadi pusat munculnya fenomena seperti IOD. IOD memiliki pengaruh besar terhadap karakteristik oseanografi seperti arah arus laut, variasi suhu permukaan laut, hingga distribusi klorofil-a.

Fenomena IOD ditandai dengan adanya anomali suhu permukaan laut: ketika bagian barat Samudra Hindia mengalami pemanasan, bagian timurnya cenderung mengalami pendinginan, atau sebaliknya. Pendinginan suhu di wilayah tertentu ini sering kali memicu upwelling, yakni proses naiknya air laut dari lapisan dalam ke permukaan karena perpindahan massa air permukaan oleh arus laut (Saji dkk., 1999). Proses ini dapat menurunkan suhu permukaan laut hingga sekitar 2°C di daerah tropis, bahkan lebih di daerah subtropis (Dahuri dkk., 1996).



Salah satu dampak signifikan oleh IOD adalah perairan selatan yang mengalami peningkatan interaksi antara lautan dan atmosfer di wilayah tropis. IOD sangat berpengaruh ketika terjadi perbedaan suhu yang mencolok antara perairan timur Afrika dan bagian timur

1.3.5 Suhu Permukaan Laut

Suhu Permukaan Laut (SPL) adalah parameter oseanografi penting yang memengaruhi sirkulasi atmosfer, cuaca, dan iklim baik regional maupun global (Alfajri dkk., 2017). Ketidakseimbangan penerimaan radiasi matahari, terutama di wilayah ekuator seperti Indonesia, memicu konveksi kuat dan pembentukan awan (Handiana dkk., 2013; Tjasyono, 2004). SPL di Indonesia juga turut membentuk pola iklim di Samudra Pasifik, Samudra Hindia, serta Benua Asia dan Australia (Purwandani, 2012).

Fluktuasi SPL berkontribusi terhadap meningkatnya curah hujan karena Indonesia berada di zona siklon tropis, wilayah bertekanan rendah yang mendorong terbentuknya awan konvektif dan cuaca ekstrem (Hartoko, 2013). Suhu juga memengaruhi kehidupan organisme laut, terutama dalam metabolisme dan reproduksi (Hutabarat & Evans, 2014). Suhu ideal bagi pertumbuhan plankton tropis berkisar antara 25–32°C; organisme eurythermal dapat hidup dalam rentang suhu luas, sedangkan stenothermal hanya pada rentang sempit (Hartoko, 2013; Wyrski, 1961).

Distribusi suhu laut bergantung pada letak lintang. Suhu tertinggi biasanya ditemukan di lintang 10°LU–10°LS karena menerima radiasi matahari paling besar (Sverdrup dkk., 1942). SPL di Indonesia umumnya berkisar 28–31°C (Nontji, 2007). Pola SPL dapat menunjukkan fenomena oseanografi seperti *upwelling* yang memicu peningkatan nutrisi, klorofil, serta penurunan suhu, dan berdampak pada distribusi fitoplankton.

Klorofil-a sebagai indikator utama produktivitas fitoplankton memiliki hubungan yang erat dengan suhu permukaan laut (SST). Umumnya, peningkatan SST menyebabkan penurunan konsentrasi klorofil-a karena suhu tinggi memperkuat stratifikasi kolom air dan menghambat proses *upwelling*, sehingga pasokan nutrisi ke permukaan berkurang dan menurunkan pertumbuhan fitoplankton. Sebaliknya, suhu yang lebih rendah dapat meningkatkan suplai nutrisi dan mendukung pertumbuhan fitoplankton. Namun, pola ini dapat bervariasi tergantung pada kondisi oseanografi dan dinamika lokal (Zhang dkk., 2020; Wu dkk., 2021).

1.3.6 Salinitas

Salinitas merupakan kadar zat terlarut (garam) dalam air laut yang diukur dalam satuan permil (‰). Bersama dengan suhu, salinitas memengaruhi berat jenis air



Salinitas umumnya berkisar antara 34–35‰. Perbedaan salinitas mengharuskan organisme beradaptasi, tetapi tidak semua spesies mampu. Organisme euryhaline, seperti ikan salmon, dapat hidup di berbagai salinitas, sedangkan organisme stenohaline, seperti kepiting, hanya bertahan di salinitas yang stabil (Tamimi dkk., 2023).

Faktor-faktor seperti penguapan, curah hujan, aliran sungai, dan sirkulasi air memengaruhi distribusi salinitas. Di wilayah pesisir, salinitas cenderung menurun saat terjadi aliran air tawar atau hujan, sementara di laut lepas, angin mencampur lapisan air hingga kedalaman 50–70 meter (Nontji, 1993). Salinitas juga memengaruhi proses reproduksi dan penyebaran ikan (Reddy, 1993). Berdasarkan kadarnya, air dibagi menjadi air tawar, air payau, dan air laut. Air payau, hasil campuran air tawar dan laut, memiliki salinitas di atas 0,5 ppt dan mengandung ion-ion utama seperti Na^+ , K^+ , Ca^{2+} , Mg^{2+} , Cl^- , SO_4^{2-} , dan HCO_3^- (Purwaningtyas dkk., 2020).

Salinitas berkaitan erat dengan suhu dan pH. Saat suhu turun, salinitas cenderung menurun karena air dingin menampung garam lebih sedikit. Air dengan salinitas tinggi biasanya memiliki pH lebih tinggi karena garam dapat menetralkan asam atau basa (Yolanda, 2023). Interaksi antara salinitas, suhu, dan pH berpengaruh besar terhadap kualitas perairan.

Menurut Pamungkas (2020), salinitas juga berperan dalam menentukan potensi tangkapan ikan, seperti cakalang, karena memengaruhi kemampuan osmoregulasi organisme laut. Oleh karena itu, salinitas dapat dijadikan indikator untuk menentukan lokasi penangkapan ikan yang potensial.

1.3.7 Pola Arus di Perairan Indonesia

Arus laut merupakan pergerakan massa air yang mengalir, yang terjadi akibat pengaruh tiupan angin, perbedaan densitas air, maupun gelombang panjang. Saat ini, arus laut telah dimanfaatkan dalam berbagai bidang yang mendukung aktivitas manusia. Namun demikian, kajian ilmiah mengenai arus laut, khususnya di wilayah perairan Indonesia, masih tergolong terbatas (Daruwedho et al., 2016). Arus laut permukaan merupakan jenis arus yang terjadi di lapisan atas perairan dan digerakkan oleh tiupan angin. Energi angin berkontribusi terhadap terbentuknya arus permukaan, dengan kecepatan arus sekitar 2% dari kecepatan angin yang bertiup. Semakin dalam lapisan air, pengaruh angin akan semakin berkurang dan umumnya tidak lagi berpengaruh pada kedalaman lebih dari 200 meter (Octavia et al., 2018).

Arus laut dapat terbentuk karena tiupan angin, perbedaan densitas, ataupun gelombang panjang yang menyebabkan massa air bergerak. Faktor-faktor yang memengaruhi dinamika arus laut antara lain arah angin, perbedaan tekanan dan



Coriolis serta arus Ekman, topografi dasar laut, arus ana *upwelling*, dan *downwelling*. Arus permukaan biasanya erapa ratus meter dari permukaan laut, bergerak secara jat dipengaruhi oleh pola angin yang dominan. Secara umum, an pergerakan massa air baik secara vertikal maupun

horizontal dalam upaya mencapai kondisi keseimbangan (Daruwedho et al., 2016).

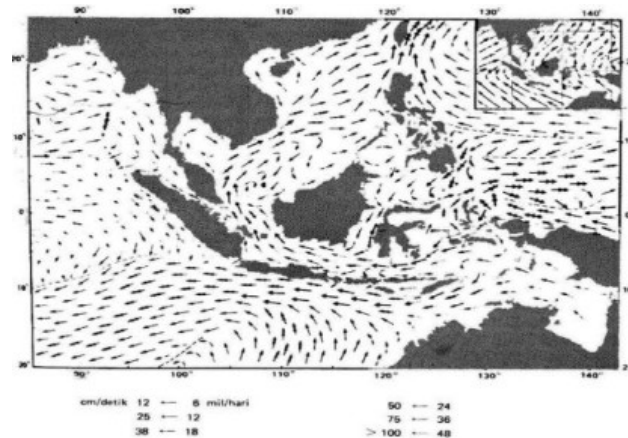
Menurut Gross (1990), arus laut dapat diklasifikasikan menjadi empat jenis berdasarkan gaya pembentuknya, yaitu:

1. Arus Ekman, yang dipicu oleh tiupan angin;
2. Arus Pasang Surut, yang terbentuk akibat gaya pasang surut;
3. Arus Termohalin, yang terjadi karena perbedaan densitas air laut;
4. Arus Geostropik, yang disebabkan oleh gradien tekanan horizontal dan pengaruh gaya Coriolis.

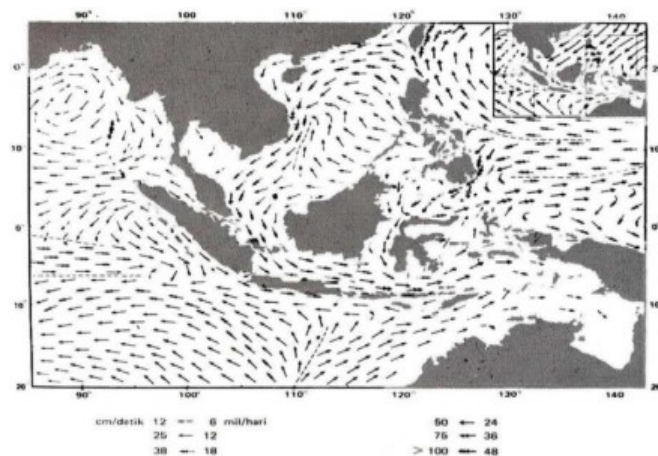
Letak geografis Asia Tenggara yang berada di antara Samudra Pasifik dan Samudra Hindia menjadikan kawasan ini sebagai jalur utama dalam sirkulasi arus laut global. Wilayah perairan seperti Laut Cina Selatan, Laut Jawa, Laut Flores, hingga Laut Banda dilalui oleh massa air yang bergerak dari timur laut ke selatan mengikuti garis khatulistiwa. Arus ini merupakan bagian dari sistem *Indonesian Throughflow* (ITF) yang mengalirkan air hangat dari Samudra Pasifik ke Samudra Hindia. Menurut Wyrki (1961), pola arus di wilayah ini berubah dua kali dalam setahun mengikuti dinamika angin musiman. Pada musim barat (Desember–Februari), arus bergerak ke arah timur dengan kecepatan yang dapat mencapai lebih dari 75 cm/s, terutama di kawasan Selat Karimata hingga Laut Flores. Pada periode ini, arus membawa massa air hangat dan turut memengaruhi kondisi oseanografi regional. Saat musim peralihan (pancaroba), arus mulai melemah dan mengalami perubahan arah, bahkan di beberapa lokasi terbentuk pusaran air (*eddies*) yang dapat memengaruhi distribusi nutrisi dan organisme laut secara signifikan.

Pada musim timur (Juni–Agustus), arah arus berbalik ke barat dan mengalir menuju Laut Cina Selatan. Pergeseran ini terjadi akibat angin monsun timur yang berasal dari Australia. Namun, di beberapa wilayah seperti sepanjang pantai utara Flores hingga Kepulauan Alor, arus pantai tetap mengalir ke timur karena dipengaruhi oleh kondisi lokal seperti kontur garis pantai dan batimetri dasar laut. Selama musim pancaroba kedua (sekitar Oktober), arah arus kembali tidak menentu; arus ke barat mulai melemah, sementara arus timur mulai berkembang hingga mencapai puncaknya pada bulan Desember. Perubahan arah arus ini tidak hanya berdampak pada sirkulasi perairan, tetapi juga berpengaruh terhadap dinamika suhu permukaan laut, pola distribusi plankton, migrasi ikan, yang merupakan daerah penangkapan. Oleh karena itu, dapat pola musiman arus sangat penting, terutama dalam daya kelautan dan perikanan. Ilustrasi mengenai perubahan arus musim-musim tersebut dapat dilihat secara visual pada Gambar 4.





Gambar 3. Pola Arus di Perairan Indonesia Bulan Juni
Sumber: Wrytki, 1961



Gambar 4. Pola Arus di Perairan Indonesia pada bulan Desember
Sumber: Wrytki, 1961

1.3.8 Linear Regression

Metode regresi adalah pendekatan statistik yang digunakan untuk meramalkan suatu nilai berdasarkan hubungan matematis antara dua jenis variabel, yakni variabel terikat (dependen/Y) dan variabel bebas (independen/X) (Khotimah & Variabel terikat merupakan hasil atau dampak yang muncul perubahan variabel bebas, sedangkan variabel bebas adalah pengaruh atau menentukan perubahan tersebut (Kusumawati nilai dari variabel bebas diketahui, maka nilai variabel terikat lakimah dkk., 2015). Dalam banyak kasus, seperti dalam variabel terikat sering kali berupa permintaan atau volume



penjualan yang dipengaruhi oleh sejumlah faktor eksternal sebagai variabel bebas (Rival dkk., 2012).

Regresi linear merupakan salah satu teknik yang digunakan dalam dunia produksi untuk memprediksi atau meramalkan karakteristik yang berkaitan dengan kualitas maupun kuantitas (Marbun et al., 2018). Metode ini digunakan karena memungkinkan perusahaan untuk mengestimasi berbagai kombinasi produk secara optimal, sehingga dapat meningkatkan efisiensi produksi dan memaksimalkan keuntungan (Indarwati et al., 2019). Persamaan regresi linear sederhana biasanya diperoleh melalui pendekatan metode kuadrat terkecil, dengan rumus sebagai berikut::

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2)} \quad (1)$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \quad (2)$$

$$y = a + b.x$$

Keterangan:

y = nilai variabel terikat (misalnya kuantitas penjualan)

x = nilai variabel bebas (misalnya waktu atau periode penjualan)

a = konstanta regresi, menunjukkan nilai y saat x = 0

b = koefisien regresi, menggambarkan besarnya perubahan nilai y untuk setiap satu satuan perubahan pada x

Dalam hal ini, y merepresentasikan jumlah penjualan, x menggambarkan periode atau bulan penjualan, a merupakan konstanta yang menunjukkan nilai y saat x bernilai nol, sedangkan b menunjukkan seberapa besar perubahan nilai y untuk setiap kenaikan satu satuan pada x.

1.3.9 Random Forest

Random Forest adalah sebuah metode yang dikembangkan oleh Breiman pada tahun 2001. Metode ini terutama digunakan untuk dua tujuan, yaitu klasifikasi dan prediksi. Secara teknis, *Random Forest* bekerja dengan memanfaatkan decision tree (pohon keputusan) sebagai dasarnya. Dengan demikian, metode ini dapat dianggap sebagai sekumpulan pohon keputusan yang bekerja bersama untuk



Proses analisisnya dimulai dengan memasukkan data melalui dan mengalirkannya hingga mencapai bagian daun (bawah).

Pada proses klasifikasi, *Random Forest* menghasilkan g-masing pohon yang terbentuk, sedangkan untuk prediksi, sh dari rata-rata nilai yang diprediksi oleh seluruh pohon

Metode *Random Forest* merupakan pengembangan dari metode *Classification and Regression Tree (CART)*, yang memanfaatkan pendekatan *bagging* (bootstrap aggregating) dan pemilihan fitur secara acak. Teknik *bagging* sendiri merupakan salah satu metode *ensemble* yang dapat meningkatkan performa algoritma klasifikasi. Menurut Rahmawati (2015), algoritma *Random Forest* bekerja dengan langkah-langkah sebagai berikut:

1. Mengambil sejumlah sampel dari dataset asli menggunakan teknik *bootstrap resampling* dengan pengembalian,
2. Membangun pohon klasifikasi dari setiap sampel bootstrap dan memilih variabel atribut secara acak untuk menentukan pemisah terbaik. Jumlah variabel yang dipilih secara acak bisa dihitung menggunakan $\frac{1}{2}\sqrt{m}$, $\sqrt{m}$, atau $2\sqrt{m}$, di mana m adalah jumlah total variabel prediktor,
3. Melakukan prediksi terhadap data berdasarkan pohon yang telah dibentuk, dan
4. Mengulangi proses tersebut hingga terbentuk sejumlah pohon klasifikasi yang diinginkan, serta melakukan proses ini sebanyak k kali.

Nilai m di sini merepresentasikan jumlah variabel prediktor yang digunakan dalam pemisahan simpul pohon. Semakin besar nilai m , maka semakin tinggi korelasi antar pohon dalam model.

Pada metode *Random Forest*, banyaknya pohon yang dibentuk menciptakan kumpulan pohon atau hutan (*forest*), dan analisis dilakukan terhadap keseluruhan hutan tersebut. Dengan data yang terdiri atas n pengamatan dan p variabel penjelas, langkah-langkah *Random Forest* menurut Breiman (2001) meliputi:

1. Melakukan pengambilan sampel acak berukuran n dari data menggunakan metode *bootstrap*,
2. Menggunakan sampel tersebut untuk membentuk pohon keputusan hingga maksimal tanpa pemangkasan. Pada setiap simpul, dilakukan pemilihan acak terhadap m variabel penjelas (dengan m jauh lebih kecil dari p), lalu dipilih pemisah terbaik dari m variabel tersebut (*random feature selection*),
3. Langkah tersebut diulang sebanyak k kali hingga terbentuk k pohon.

Untuk prediksi, hasil dari seluruh pohon dikombinasikan (digabungkan) menjadi satu, di mana dalam kasus klasifikasi digunakan metode *majority vote* atau suara terbanyak untuk menentukan kelas akhir dari suatu pengamatan.



ss Validation

ss validation merupakan pendekatan yang efektif dalam in untuk mengintegrasikan proses pemilihan fitur (*feature* nyetelan parameter (*parameter tuning*) guna membangun

model prediktif yang optimal. Dalam teknik ini, himpunan data dibagi menjadi k kelompok (lipatan atau *folds*) yang berukuran sama. Setiap iterasi menggunakan $k-1$ lipatan sebagai data pelatihan (*training set*) dan satu lipatan sisanya sebagai data pengujian (*testing set*). Proses ini diulang sebanyak k kali, sehingga setiap lipatan menjadi data pengujian tepat satu kali. Teknik ini dinilai unggul karena memaksimalkan penggunaan data secara efisien untuk pelatihan dan evaluasi, sekaligus mengurangi kemungkinan overfitting (Parvande dkk, 2020).

Dalam penerapannya, *k-fold cross validation* dapat dikombinasikan dengan validasi bersarang (*nested validation*) pada proses pelatihan bagian dalam (*inner loop*) untuk melakukan seleksi fitur dan penyetelan parameter model secara lebih akurat. Validasi bersarang ini mencegah terjadinya kebocoran informasi antar lipatan luar (*outer folds*), yang jika tidak dikendalikan dapat menyebabkan bias dalam proses pemilihan fitur dan estimasi performa model. Oleh karena itu, pendekatan ini sangat penting untuk memastikan bahwa proses pelatihan dan evaluasi benar-benar independen, khususnya ketika model akan diterapkan pada data baru yang belum pernah dilihat sebelumnya (Parvande dkk, 2020).

Pemilihan nilai k sangat menentukan stabilitas dan akurasi model yang dibangun. Jika k terlalu besar (misalnya mendekati jumlah total data), hasil validasi cenderung memiliki bias rendah namun varians tinggi, yang dapat memicu overfitting. Sebaliknya, jika k terlalu kecil, hasil validasi lebih bias karena data pelatihan lebih sedikit. Oleh karena itu, dalam praktik umum disarankan menggunakan nilai k sebesar 5 atau 10, yang dianggap seimbang antara bias dan varians (Kuhn & Johnson, 2013).

Setelah data dibagi menjadi k bagian, sebanyak $k-1$ bagian digunakan untuk melatih model, dan satu bagian sisanya digunakan untuk menguji performa model. Evaluasi dilakukan dengan menghitung *root mean square error* (RMSE) pada setiap iterasi. RMSE dihitung menggunakan rumus:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$

$\hat{y}_i$ adalah nilai prediksi dan y_i adalah nilai aktual pada pengamatan ke- i . Proses ini diulang sebanyak k kali sehingga menghasilkan $RMSE_1, RMSE_2, \dots, RMSE_k$. Untuk mengevaluasi performa keseluruhan, digunakan nilai rata-rata RMSE dari seluruh iterasi, dirumuskan sebagai:

$$CV_{(k)} = \frac{1}{k} \sum_{i=1}^k RMSE_i$$



Model terbaik adalah model yang memiliki nilai rata-rata RMSE terkecil, karena menunjukkan bahwa kesalahan prediksi berada pada tingkat yang paling rendah (Zhang, 2019). Nilai RMSE yang kecil menandakan bahwa prediksi model mendekati nilai aktual, sehingga tingkat deviasi antara hasil prediksi dan observasi sesungguhnya relatif rendah. Hal ini sangat penting dalam konteks pemodelan prediktif, terutama ketika data digunakan untuk pengambilan keputusan atau proyeksi di masa mendatang. Selain itu, RMSE memiliki satuan yang sama dengan variabel target, sehingga interpretasinya menjadi lebih intuitif dibandingkan metrik kesalahan lain seperti MSE atau MAE. Oleh karena itu, dalam proses validasi model, rata-rata RMSE yang rendah tidak hanya merefleksikan performa model yang baik secara statistik, tetapi juga mengindikasikan keandalan model dalam penerapan praktis pada data yang belum pernah dilihat sebelumnya.

1.3.11 Bayesian Optimization

Bayesian *optimization* merupakan strategi pencarian parameter optimal yang dirancang untuk menyelesaikan permasalahan fungsi objektif yang tidak diketahui bentuk matematisnya (*black-box*) dan mahal untuk dievaluasi. Metode ini bekerja dengan membuat model perantara (*surrogate model*), seperti *Gaussian Process Regression*, yang bertugas memperkirakan nilai fungsi berdasarkan data observasi yang telah dikumpulkan. Kemudian, titik parameter yang akan diuji selanjutnya ditentukan melalui *acquisition function* yang menimbang antara pencarian wilayah baru dan pendalaman pada area yang menjanjikan (Frazier, 2018).

Salah satu keunggulan dari metode ini adalah efisiensinya dalam melakukan tuning hiperparameter pada model pembelajaran mesin, khususnya ketika evaluasi model membutuhkan waktu lama dan daya komputasi yang tinggi, seperti dalam *cross-validation*. Selain itu, pendekatan ini cukup andal dalam kondisi data yang mengandung noise dan ideal untuk ruang parameter yang tidak terlalu besar (Snoek dkk., 2012).

Dalam konteks penelitian ini, *Bayesian optimization* digunakan untuk memperoleh konfigurasi hiperparameter terbaik pada algoritma *Random Forest* guna meningkatkan akurasi prediksi terhadap jumlah hasil tangkapan ikan cakalang. Proses optimasi difokuskan pada sejumlah parameter utama seperti jumlah pohon (*n_estimators*), jumlah minimum sampel di daun akhir, serta yang dipertimbangkan pada setiap pemisahan. Secara acak sebanyak 20 kali evaluasi terhadap fungsi objektif, di mana menghasilkan kombinasi baru yang dinilai berdasarkan nilai error (RMSE) melalui pendekatan *k-fold cross-validation*.



Selama proses tersebut berlangsung, *Bayesian optimization* secara progresif mengarahkan pencarian ke wilayah parameter yang memberikan kinerja prediksi lebih baik. Model perantara membantu memetakan area-area yang menjanjikan di ruang parameter, sementara mekanisme *acquisition function* memastikan titik-titik evaluasi berikutnya mempertimbangkan keseimbangan antara mencoba hal baru dan memperdalam yang sudah baik. Setelah seluruh siklus evaluasi selesai, diperoleh kombinasi parameter paling optimal yang kemudian diterapkan pada tahap akhir pemodelan *Random Forest*. Konfigurasi inilah yang digunakan untuk menghasilkan prediksi tangkapan ikan cakalang dengan akurasi yang lebih tinggi.

