**Faktor yang Memengaruhi Kehidupan Organisme di Laut Berdasarkan Data CalCOFI: *Over 60 Years of Oceanographic Data***

Kelangsungan hidup orgamisme di bawah laut sangatlah bergantung terhadap perubahan iklim dan temperatur. Temperatur atau suhu merupakan salah satu faktor yang penting dalam mengatur proses kehidupan dan penyebaran organisme, karena suhu mempengaruhi baik aktivitas maupun perkembangbiakan dari organisme tersebut. Dalam oseanografi dikenal dua istilah untuk menentukan temperatur air laut yaitu temperatur insitu dan temperatur potensial. Semakin besar aktivitas (energi), semakin tinggi pula temperaturnya. Temperatur menunjukkan kandungan energi panas. Temperatur air laut di permukaan ditentukan oleh adanya pemanasan (*heating*) di daerah tropis dan pendinginan (*cooling*) di daerah lintang tinggi. Kisaran harga temperatur di laut adalah -2o s.d. 35oC. Selain suhu, kehidupan di bawah laut juga dipengaruhi oleh beberapa faktor yang menyebabkan suhu potensial dalam laut dapat dinyatakan layak untuk kehidupan organisme disana. Suhu poCalCOFI merupakan salah satu badan yang berfokus pada studi tentang lingkungan laut di lepas pantai California, pengelolaan sumber daya kehidupannya, dan memantau indikator El Nino dan perubahan iklim. Data yang dikumpulkan pada kedalaman hingga 500 m meliputi: suhu, salinitas, oksigen, fosfat, silikat, nitrat dan nitrit, klorofil, transmissometer, PAR, produktivitas primer C14, keanekaragaman hayati fitoplankton, biomassa zooplankton, dan keanekaragaman hayati zooplankton. Berdasarkan data yang telah dikumpulkan, didapatkan data yang kurang lengkap atau terdapat *missing value* sehingga dibutuhkan *preprocessing* data sebelum mengolahnya lebih lanjut. Pada penelitian kali ini digunakan beberapa variabel yang memengaruhi kehidupan organisme di laut antara lain temperatur potensial, suhu, kedalaman laut, salinitas, densitas air laut, pasang surut air laut, kadar oksigen dalam air laut, intensitas oksigen dalam air, volume spesifik anomali, dan tekanan dalam air laut. Data yang digunakan adalah data dengan satu variabel independen yaitu temperatur potensial dan 9 variabel dependen lainnya serta jumlah pengamatan yang digunakan sebanyak 5000 pengamatan.

1. *Missing Value*

Kelengkapan data sangatlah penting untuk mendapat hasil analisis data yang akurat. Data yang tidak lengkap atau memiliki *missing value* harus diatasi terlebih dahulu dengan proses *preprocessing* data. Berdasarkan hasil deteksi *missing value* didapatkan variabel yang memiliki data *missing* atau kurang lengkap adalah sebagai berikut.

|  |  |
| --- | --- |
| **Variabel** | ***Missing Value*** |
| Temperatur Potensial | 225 |
| Temperatur | 20 |
| Salinitas | 152 |
| Densitas potensial | 281 |
| volume spesifik anomali | 281 |
| Pasang surut air laut | 214 |
| Kadar oksigen dalam air laut | 2190 |
| Persentase intensitas oksigen dalam air | 2310 |

Hasil deteksi missing value dapat diatasi dengan menggunakan mean pada masing-masing variabel jika data *missing* tidak lebih dari 50% dan memiliki korelasi kuat dengan variabel independen yang diamati. Dengan menggunakan statistika deskriptif didapatkan nilai mean sebagai berikut.

|  |  |
| --- | --- |
| **Variabel** | **Nilai Rata-rata** |
| Temperatur Potensial | 9,004084 |
| Temperatur | 8,974335 |
| Salinitas | 33,828065 |
| Densitas potensial | 26,120341 |
| volume spesifik anomali | 192,307163 |
| Pasang surut air laut | 0,587048 |
| Kadar oksigen dalam air laut | 3,105388 |
| Persentase intensitas oksigen dalam air | 50,474572 |

Setelah diatasi dengan nilai rata-rata dan data sudah lengkap maka diperoleh hasil model regresi sebagai berikut.

R\_POTEMP = - 425 + 0.273 R\_TEMP + 0.0843 R\_DEPTH + 2.49 R\_SALINITY

+ 12.1 R\_SIGMA + 0.166 R\_SVA - 1.96 R\_DYNHT - 0.0833 R\_PRES

- 0.382 R\_O2 + 0.0237 R\_O2Sat

Berdasarkan model di atas didapatkan nilai p-value tiap variabel sebesar 0,0000 pada tingkat kepercayaan 95% dan alfa 0,05 yang artinya tiap variabel berpengaruh secara signifikan terhadap model.

1. *Outlier*

Selain kelengkapan data, hal yang harus diperhatikan sebelum mengolah data adalah data *outlier* atau data ekstrim sehingga dapat mempengaruhi hasil uji asumsi atau bahkan dalam pengambilan keputusan yang salah. Oleh karena itu, sebelum melakukan analisis dibutuhkan deteksi *outlier* dan diatasi terlebih dahulu. Berdasarkan hasil deteksi *outlier* didapatkan 3 pengamatan yang terdeteksi sebagai *outlier* dan dihapuskan dengan asumsi telah diketahui penyebabnya. Setelah ketiga pengamatan penyebab *outlier* tersebut dihapuskan, didapatkan hasil model regresi sebagai berikut.

R\_POTEMP = - 425 + 0.273 R\_TEMP + 0.0859 R\_DEPTH + 2.49 R\_SALINITY

+ 12.1 R\_SIGMA + 0.167 R\_SVA - 1.97 R\_DYNHT - 0.0850 R\_PRES

- 0.382 R\_O2 + 0.0237 R\_O2Sat

1. *Missing Value* dan Transformasi

Transformasi data adalah upaya untuk yang dilakukan dengan tujuan utama untuk mengubah skala data pengukuran data asli menjadi bentuk lain sehingga data dapat memenuhi asumsi-asumsi yang mendasari analisis ragam. Dengan menggunakan data yang sudah lengkap kemudian dilakukan transformasi *softmax* didapatkan model regresi sebagai berikut.

R\_POTEMP = 8.70 + 894140 R\_TEMP + 407891 R\_DEPTH + 165207 R\_SALINITY

- 632813 R\_SIGMA + 90535 R\_SVA - 318512 R\_PRES - 659550 R\_O2

+ 54267 R\_O2Sat

1. *Missing Value*, *Outlier*, Transformasi

Setelah mengatasi *missing value* dan *outlier*, data dilakukan transformasi data untuk mengubah skala data dengan menggunakan metode *softmax* sehingga didapatkan hasil model regresi sebagai berikut.

R\_POTEMP = 929 + 966321 R\_TEMP + 243474 R\_DEPTH + 178454 R\_SALINITY

- 624495 R\_SIGMA + 81963 R\_SVA - 4717834 R\_DYNHT – 161494

R\_PRES - 615723 R\_O2 + 49090 R\_O2Sat

Berdasarkan hasil *preprocessing* data diperoleh empat model regresi berbeda dengan nilai

*R-square* sebagai berikut.

|  |  |  |  |
| --- | --- | --- | --- |
| **Metode** | **standar deviasi** | **R-sq** | **R-sq (adj)** |
| Missing Value | 0,418177 | 98.8% | 98.8% |
| Missing Value dan Outlier | 0,418250 | 98.8% | 98.8% |
| Missing Value dan Transformasi | 1,16542 | 90.6% | 90.6% |
| Missing Value, Outlier, dan Transformasi | 1,14133 | 91.0% | 90.9% |

Hasil *R-square* maksimum dapat membantu dalam menentukan manakah model terbaik. Berdasarkan tabel di atas dapat diketahui bahwa data yang telah diatasi *missing value* dan *outlier* merupakan model terbaik dengan *R-square* maksimum yaitu 98,8% dan standar deviasi minimum yaitu 0,418250. Pada tabel tersebut menunjukkan bahwa *outlier* mampu mempengaruhi penyebaran data yang dapat dilihat dari nilai standar deviasinya, sedangkan transformasi dapat meningkatkan ragam pada data dan mempengaruhi *R-square.* Oleh karena itu, dapat disimpulkan model terbaik dari keempat metode ini adalah sebagai berikut.

R\_POTEMP = - 425 + 0.273 R\_TEMP + 0.0859 R\_DEPTH + 2.49 R\_SALINITY

+ 12.1 R\_SIGMA + 0.167 R\_SVA - 1.97 R\_DYNHT - 0.0850 R\_PRES

- 0.382 R\_O2 + 0.0237 R\_O2Sat