

1. Diagram Taylora

W celu umożliwienia wieloaspektowej oceny jakości symulacji często stosuje się graficzne przedstawienie rezultatów wydajności modeli z wykorzystaniem tzw. diagramów Taylora (2001). Diagram ten pozwala na uwzględnienie na jednym wykresie 3 miar charakteryzujących jakość modelu.

1. Korelacja Pearsona (R) - określa jakość modelu do odtwarzania czasowej zmienności analizowanego elementu. Wielkość korelacji jest ilorazem kowariancji i iloczynu odchyłeń standardowych analizowanych zmiennych (równanie 1.1).
2. Średni błąd kwadratowy (RMSE) - określa statystyczną "odległość" danych modelowych względem obserwacji. Pozwala na wyrażenie wartości błędu w tych samych jednostkach w jakich jest on mierzony (Wilks 2008) (równanie 1.2).
3. Odchylenie standardowe - umożliwia porównanie amplitudy wartości serii referencyjnej (równanie 1.3) i modelowej (równanie 1.4).

$$R = \frac{\frac{1}{N} \sum_{n=1}^N (f_n - f)(r_n - r)}{\sigma_f \sigma_r} \quad (1.1)$$

$$RMSE = \frac{1}{N} \sum_{n=1}^N [(f_n - f) - (r_n - r)]^2 \quad (1.2)$$

$$\sigma_r^2 = \frac{1}{N} \sum_{n=1}^N (r_n - r)^2 \quad (1.3)$$

$$\sigma_f^2 = \frac{1}{N} \sum_{n=1}^N (f_n - f)^2 \quad (1.4)$$

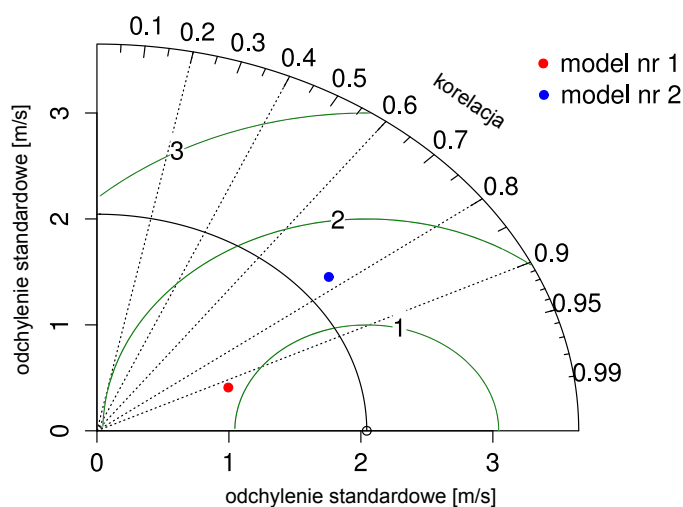
gdzie:

f - seria danych modelowych

r - seria danych referencyjnych

σ_f, σ_r - odchylenie standardowe serii modelowej (σ_f) i referencyjnej (σ_r)

Ze względu na różne odmiany diagramów Taylora poniżej omówiono jedną z najbardziej popularnych jego wersji, którą zastosujemy w dalszej części ćwiczenia. Porównanie dwóch losowo wygenerowanych symulacji modelowych względem serii obserwacyjnej przedstawiono na rys. 1.1. Na osiach rzędnych i odciętych opisane wartości dotyczą parametru odchylenia standardowego (σ), natomiast łącząca je krzywa (oznaczona ciągłą czarną linią) zakreśla wartość tego parametru dla serii referencyjnej. Oznaczone kolorem zielonym linie o kształcie zbliżonym do półokręgów wskazują wartości średniego błędu kwadratowego (RMSE), natomiast przerywane linie koloru czarnego, biegnące od początku diagramu do jego prawego górnego ograniczenia wytyczają wartości współczynników korelacji (r), opisanych na krawędzi wykresu. Na osi odciętych oznaczono za pomocą białego punktu z czarnym obramowaniem wartość serii referencyjnej.



Rysunek 1.1. Przykład zastosowania diagramu Taylora do opisu wyników symulacji modelowych

Przykładowe statystyki testowe modelu nr 1 przedstawione na czerwono na rys. 1.1 wynoszą zatem: $\sigma = 1,1$ m/s, RMSE = 1,15 m/s, $r = 0,91$, natomiast dla modelu nr 2 (kolor niebieski) są to wartości: $\sigma = 2,25$ m/s, RMSE = 1,6 m/s oraz $r = 0,78$.

Diagram Taylora – ćwiczenie

1. W pliku taylor.xls zawarto dane prognozy pogody z 8 różnych modeli globalnych dla 2 stacji meteorologicznych (kolumna 'station'), z wyprzedzeniem T0+48h.
2. Należy zaznaczyć na diagramie Taylora „położenie” oznaczające wydajność prognostyczną każdego z modeli, dla obu analizowanych stacji (w sumie powinno być 16 punktów).
3. Stwórz legendę do zaznaczonych punktów i dołącz ją do opracowania.
4. Który z analizowanych modeli jest Twoim zdaniem najlepszy, a który najgorszy?

1.1. Jak to zrobić w R?

W programie R rysowanie diagramów Taylora jest możliwe dzięki funkcji `taylor.diagram` z biblioteki `plotrix` (na większości komputerów najpierw trzeba ją zainstalować).

1. W pierwszej kolejności należy zapisać pobrany zbiór danych z formatu xls na csv, a następnie wczytać do programu R za pomocą polecenia `read.csv("nazwanaszegopliku.csv")`. Jeśli nasz nowy zbiór w R nazwiemy jako "dane" wówczas polecenie będzie miało postać:

```
dane=read.csv("nazwanaszegopliku.csv")
```

Należy pamiętać o wcześniejszym wskazaniu katalogu roboczego, w którym znajduje się nasz plik (górne menu: "session" → set working directory → choose directory)

2. Po aktywacji paczki `plotrix` możemy korzystać z funkcji `taylor.diagram`. Po nazwie funkcji należy podać zbiór danych traktowanych jako serię referencyjną oraz serię danych modelowych.
3. Dodanie każdego kolejnego punktu do diagramu Taylora jest możliwe poprzez wywołanie wcześniej opisywanej funkcji z opcją `add=TRUE`
4. Warto zadbać o odpowiednie definiowanie kolorów (opcja: `col="nazwa koloru po ang."`) lub wielkości dodawanych punktów (opcja: `cex=wartość`) lub ich kształtu (opcja: `pch=liczba całkowita od 1 do 21`)