

Uniwersytet Warszawski
Wydział Matematyki, Informatyki i Mechaniki

Marcin Osowski

Nr albumu: 292682

Hyper Thermic

Mapy termiki szybowcowej

Praca licencjacka
na kierunku MATEMATYKA

Praca wykonana pod kierunkiem
dr Marcin Szczuka

Grudzień 2014

Oświadczenie kierującego pracą

Potwierdzam, że niniejsza praca została przygotowana pod moim kierunkiem i kwalifikuje się do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Data

Podpis kierującego pracą

Oświadczenie autora (autorów) pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data

Podpis autora (autorów) pracy

Streszczenie

W pracy przedstawiono metody oraz wyniki analizy ponad 140.000 zapisów lotów szybowcowych dostępnych publicznie w Internecie. Celem analizy było stworzenie mapy obrazującej wpływ terenu na jakość (prędkość pionową) napotykanym wznoszeń termicznych, wykorzystywanych przez pilotów szybowcowych jako podstawowe źródło energii. Do wstępnego przetworzenia zapisów i wyodrębnienia informacji o kominach termicznych użyto łańcuchów Markowa, natomiast do zbudowania map kominów termicznych użyto metody k-NN oraz dwuwymiarowych wielomianów interpolacyjnych Hermite'a. Wyniki eksperymentalne zastosowanych modeli predykcyjnych pokazują, że modele te wyłącznie na podstawie lokalizacji komina termicznego tłumaczą 7-8% wariancji prędkości pionowej. Ponadto wyniki modeli okazują się być zgodne z mapami obszarów tworzonymi ręcznie przez pilotów z olbrzymim doświadczeniem w wykonywaniu lotów na tych obszarach.

Słowa kluczowe

mapa kominów termicznych, liniowe zadanie najmniejszych kwadratów, k-NN, wielomiany interpolacyjne Hermite'a

Dziedzina pracy (kody wg programu Socrates-Erasmus)

11.1 Matematyka

Klasyfikacja tematyczna

62 Statistics

62F Parametric inference

62F10 Point estimation

Tytuł pracy w języku angielskim

Hyper Thermic: Thermal maps for gliders

Spis treści

Wprowadzenie	9
1. Wznoszenia termiczne oraz ich wykorzystanie	11
1.1. O wznoszeniach termicznych	11
1.1.1. Termika wypracowana	11
1.1.2. Termika naniesiona	12
1.2. Wykorzystanie w szybownictwie	13
1.3. Inne rozwiązania zagadnienia map szybowcowych	13
1.3.1. ParaglidingNet/Ikarus	14
1.3.2. TherMap	14
1.3.3. Track2Thermic	15
1.3.4. GlidingHotspots.eu	16
1.3.5. Mapy Janusza Centki	16
2. Mapy	19
2.1. Definicja mapy	19
2.2. Wielomiany interpolacyjne Hermite’a	20
2.2.1. Zastosowanie do budowania mapy	21
2.2.2. Metoda LSQR	22
2.2.3. Przeuczanie	22
2.3. k-NN	23
3. Dane	25
3.1. Rejestratory, format IGC	25
3.2. Zestawienie źródeł danych	26
3.3. Wykrywanie kominów termicznych	26
3.4. Ogólne statystyki danych	27
4. Wyniki eksperymentalne	31
4.1. Pomiar skuteczności klasyfikatorów	31
4.2. Mapy termiki szybowcowej	32
5. Podsumowanie i perspektywy dalszych prac	37
Bibliografia	39

Spis rysunków

1.1.	Krażenie w kominie, SeeYou	13
1.2.	Esowanie, SeeYou	13
1.3.	Drogi powietrzne, ParaglidingNet	14
1.4.	Mapa, ParaglidingNet	14
1.5.	TherMap	15
1.6.	ParaglidingNet versus TherMap	15
1.7.	Track2Thermic	16
1.8.	GlidingHotspots.eu, mapa	17
1.9.	Mapa Janusza Centki	17
1.10.	Przetworzona mapa Janusza Centki	18
2.1.	Bazowe wielomiany Hermite’a	20
3.1.	Estymator gęstości prędkości pionowej	28
3.2.	Dystrybuanta prędkości pionowej	28
3.3.	Średnia prędkość komina w ciągu dnia	29
3.4.	Średnia prędkość komina vs odległość	29
3.5.	Zlogarytmowany historgam czasu krążenia	30
3.6.	Zlogarytmowany histogram prędkości i czasu	30
4.1.	Mapa k-NN dla Polski	33
4.2.	Mapa Hermite dla Polski	34
4.3.	Porównanie z mapami Janusza Centki	35

Spis tabel

1.1. Pochłanianie promieniowania	11
3.1. Serwisy z plikami ICG	26
4.1. Wyniki klasyfikatora Hermite	31
4.2. Wyniki klasyfikatora k-NN	32

Wprowadzenie

Podstawowym zadaniem pilota szybowcowego podczas lotu jest wyszukiwanie najkorzystniejszych wznoszeń, czyli pionowych ruchów powietrza w atmosferze, które, pomijając start, są jedynym źródłem energii szybowca. Zdecydowaną większość wznoszeń można podzielić na trzy rodzaje, nazywane żagle, falą górską oraz termiką. W pracy tej będę się skupiał wyłącznie na ostatnim rodzaju, czyli na wznoszeniach termicznych, jako najbardziej powszechnych.

Aby odnaleźć możliwie najlepsze wznoszenia pilot wykorzystuje przede wszystkim: odczyty z pokładowych przyrządów, wiedzę na temat obserwowalnej części obecnego stanu atmosfery (chmury, wiatr), zachowanie innych statków powietrznych i ptaków, wiedzę z prognozy pogody, oraz charakter i kształt terenu.

Celem niniejszej pracy jest wykorzystanie metod numerycznych do zbudowania mapy wpływu terenu nad którym odbywa się lot na jakość napotykaných wznoszeń termicznych, przydatnej na przykład podczas planowania trasy lotu. W tym celu wykorzystano dostępne publicznie w Internecie zapisy lotów szybowców z wielu różnych zawodów szybowcowych z lat 2003 – 2013, łącznie ponad 140.000 zapisów.

Uzyskane wyniki pokazują, że sama informacja o lokalizacji komina termicznego jest w stanie wytłumaczyć 7-8% wariancji prędkości pionowej komina termicznego. Uzyskane wyniki są zgodne z mapami tworzonymi ręcznie przez doświadczonych pilotów na podstawie ich osobistego, kilkudziesięcioletniego doświadczenia w wykonywaniu lotów szybowcowych.

Praca została zorganizowana w następujący sposób: rozdział 1 zawiera wprowadzenie do zagadnienia prądów termicznych atmosfery, opis ich wykorzystania w szybownictwie, oraz spis istniejących rozwiązań problemu map termicznych. Rozdział 2 zawiera precyzyjny opis problemu optymalnej mapy, oraz dwie metody użyte do budowania map – dwuwymiarowe wielomiany Hermite’a, oraz metodę k-NN. Rozdział 3 zawiera informacje o źródłach danych, opis metod użytych do wstępnego przetworzenia danych, oraz krótki opis ogólnych statystycznych właściwości danych. Rozdział 4 zawiera wyniki eksperymentalne uzyskane przez zastosowanie opisanych metod budowania map na zebranych danych. Rozdział 5 zawiera podsumowanie uzyskanych wyników oraz perspektywy na dalsze prace w kierunku map termiki szybowcowej.

Rozdział 1

Wznoszenia termiczne oraz ich wykorzystanie

1.1. O wznoszeniach termicznych

W najniższej warstwie atmosfery Ziemi, troposferze, w sposób naturalny powstają pionowe, konwekcyjne prądy powietrzne wzbudzone różnicami temperatury powietrza. Prądy te są wykorzystywane przez pilotów szybowcowych jako naturalne źródło energii, pozwalając na wykonywanie lotów na trasach o długościach setek kilometrów, czasami nawet przekraczających tysiąc kilometrów. Zjawisko powstawania takich pionowych ruchów powietrza nazywane jest zwyczajowo *termiką*, a same prądy *kominami termicznymi* (ang. *thermals*).

Prądy termiczne mają bardzo duże znaczenie dla kształtowania pogody Ziemi - są istotną i obszerną częścią meteorologii. W tym opracowaniu skoncentruję się wyłącznie na istotnej dla zagadnienia części tego zjawiska. Poniższy opis bazuje na [1] oraz [2].

1.1.1. Termika wypracowana

Z termiką wypracowaną mamy do czynienia gdy Słońce intensywnie ogrzewa powierzchnię Ziemi, w konsekwencji czego ogrzewa się również warstwa powietrza bezpośrednio nad terenem. Powietrze cieplejsze ma mniejszą gęstość, i przy dostatecznej różnicy temperatur rozpocznie ruch konwekcyjny w kierunku wyższych partii atmosfery.

Warto tutaj zaznaczyć, że różne rodzaje terenu absorbują promieniowanie słoneczne w różnym zakresie. Parczewski w [1] podaje następujące dane:

Tabela 1.1: Stopień, w jakim różne powierzchnie pochłaniają promieniowanie

Rodzaj powierzchni	Ilość pochłanianego promieniowania
Pokrywa śnieżna	20-30 %
Woda pod kątem 5.5°	29 %
Sucha gleba	86 %
Łąki	67-90 %
Kamieniste podłoże	85-88 %
Wilgotna gleba	91-92 %
Woda pod kątem 47°	98 %

Ponadto na ilość energii słonecznej padającej na podłoże duży wpływ ma kąt padania promieni. Zgodnie z prawem Lamberta moc promieniowania padającego na powierzchnię jest

proporcjonalna do kosinusa kąta pomiędzy normalną tej powierzchni, a kierunkiem do źródła światła. Na obszarze Polski, w okresie letnim, promienie Słońca w zenicie padają pod kątem od około 40° do około 60° , w zależności od szerokości geograficznej oraz dnia w roku. Wartość ta może być dalej zmodyfikowana lokalnym kształtem terenu. W szczególności górskie zbocza od strony południowej są dodatkowo uprzywilejowane otrzymując więcej energii.

I tak na przykład podczas zenitu, gdy Słońce oświetla ziemię pod kątem 50° , południowe zbocze o nachyleniu 30° otrzymuje około 46% więcej energii słonecznej niż podłoże płaskie o tej samej powierzchni.

Ostatnim czynnikiem wpływającym na ogrzewanie powietrza nad terenem jest pojemność cieplna samego terenu. Tereny o dużej pojemności cieplnej będą wolniej reagowały na promieniowanie słoneczne, ale będą też w stanie dłużej ogrzewać powietrze gdy promieniowanie osłabnie. Przykładami takich terenów są: zbiorniki wodne, zielona roślinność, wilgotny piasek czy wilgotna gleba. Symetrycznie będą się zachowywały tereny o małej pojemności cieplnej – szybko staną się aktywne, ale też szybko przestaną ogrzewać powietrze gdy zniknie lub osłabnie promieniowanie słoneczne. Do tej grupy należy zaliczyć podłoża takie jak: suchy piasek, kamienie, suchą zaoraną glebę czy łąny dojrzewających zbóż.

Komin termiki wypracowanej jest inicjowany lokalnym impulsem: w pewnym momencie bąbel ciepłego powietrza odrywa się od podłoża i pociąga za sobą cały strumień ciepłego powietrza zalegającego nad terenem. Przykładem takiego impulsu jest kontrast termiczny – sąsiadujące powierzchnie, nad którymi powietrze jest podgrzewane w różnym tempie. Weźmy granicę pomiędzy jeziorem a obszerną plażą podczas dużego nasłonecznienia. Jezioro co prawda bardzo dobrze pochłania energię słoneczną (tabela 1.1), ale ma też znaczną pojemność cieplną. W skutek tego powietrze nad jeziorem będzie chłodniejsze i będzie wypierało ku gorze sąsiadujące powietrze znad plaży.

Po inicjacji komina termicznego może on pozostać „przyklejony” do konkretnego miejsca terenowego, ale może też zacząć się przesuwać. Dzieje się tak na przykład gdy prąd konwekcyjny osiągnie poziom kondensacji, i powstanie chmura typu *Cumulus*. Skraplanie pary wodnej jest zjawiskiem produkującym znaczną ilość energii i wzmacniającym zjawiska konwekcyjne, dlatego przy udziale wiatru chmura może zostać przesunięta, pociągając za sobą komin termiczny. Taki komin przestaje być powiązany z początkowym impulsem, ale nadal jego źródło energii to ogrzane powietrze zalegające bezpośrednio nad terenem.

1.1.2. Termika naniesiona

Termika naniesiona jest zjawiskiem dynamicznym, to znaczy ściśle powiązany z wiatrem. Typowo ten rodzaj termiki objawia się przy wietrze o prędkości przekraczającej 5 m/s oraz napływie chłodnych mas powietrza arktycznego lub polarnego. Prądy termiki naniesionej są wyzwalane przez ruch oraz zawirowania powietrza powodowane wiatrem, w szczególności sprzyjające są tu:

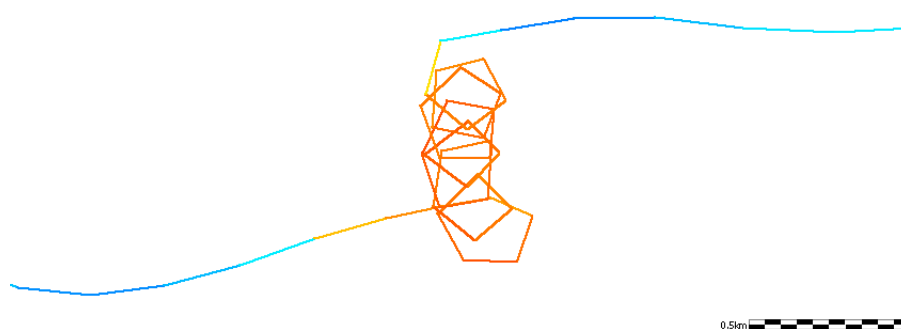
- nierówności terenowe: wzgórza, budynki, ściany lasu
- nierównomierne tarcie o podłoże na granicy podłoża szorstkiego i gładkiego
- ruchy falowe
- powierzchnie rozległych obszarów wodnych
- znaczna prędkość wiatru powiązana z tarcieniem o podłoże

Tak jak w przypadku termiki wypracowanej możemy wyróżnić ogniska termiczne, czyli miejsca, w których prawdopodobieństwo powstania komina termiki naniesionej jest większe.

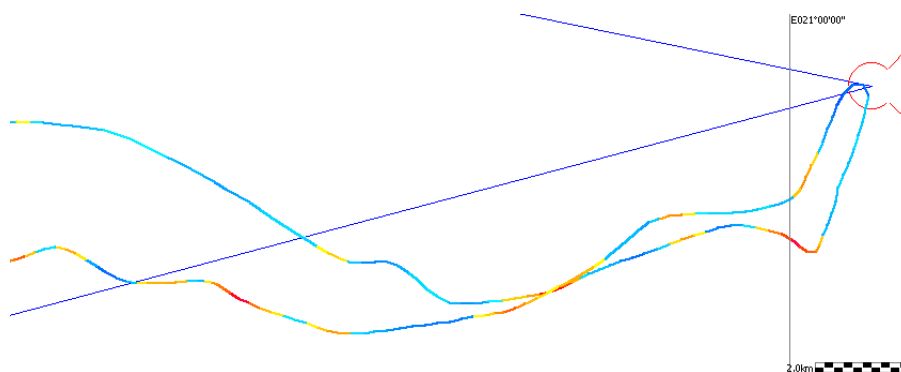
1.2. Wykorzystanie w szybownictwie

Kominy termiczne przyjmują postać mniej lub bardziej regularnych słupów powietrza, o przekrojach w przybliżeniu w kształcie koła (lub elipsy). Prędkość pionowa typowo jest maksymalna w centrum, i osiąga zazwyczaj wartości od 0 m/s do 10 m/s (pomijając warunki burzowe). Na terenie Polski najczęściej spotykane kominy mają maksymalne prędkości pionowe z zakresu od 1 m/s do 5 m/s .

Piloci szybowcowi wykorzystują wznoszenia termiczne albo poprzez krążenie wewnątrz kominów termicznych (rys 1.1), albo poprzez *esowanie* pod połączonymi chmurami Cumulus (rys 1.2). Esowanie to zabieg polegający na manipulacji prędkością szybowca, powiązanej z niewielkimi odchyleniami od zamierzonego kursu, w celu wykorzystania wznoszeń termicznych bez potrzeby wykonywania krążenia. Takie ścieżki połączonych wznoszeń termicznych są zwyczajowo nazywane *szlakami* Cumulusów (ang. *cloud street*).



Rysunek 1.1: Zapis z rejestratora – pilot krąży w kominie termicznym; program SeeYou



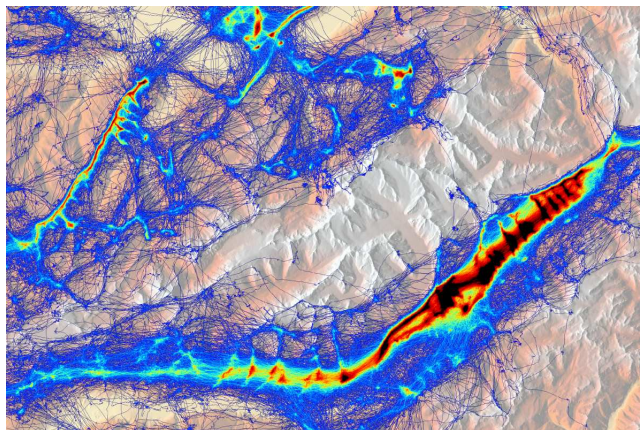
Rysunek 1.2: Zapis z rejestratora – pilot esuje, zalicza punkt, po czym wraca; program SeeYou

1.3. Inne rozwiązania zagadnienia map szybowcowych

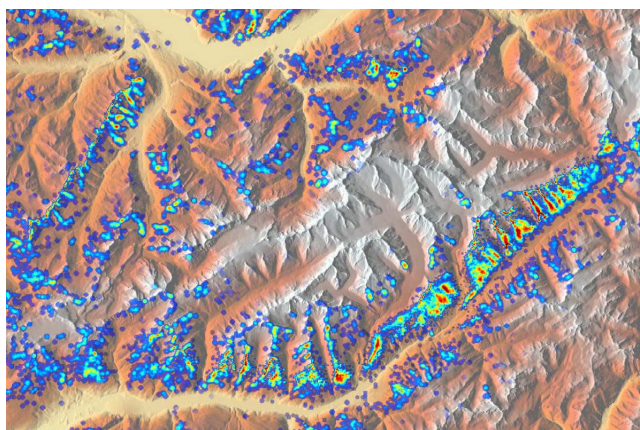
Pomysł tworzenia map termiki szybowcowej jest prawdopodobnie niewiele młodszy niż samo szybownictwo. W niniejszej sekcji opiszę znane mi podejścia tworzenia map szybowcowych, a także, ze względu na istotną zbieżność interesów, map termicznych dla użytkowników paralotni i lotni.

1.3.1. ParaglidingNet/Ikarus

ParaglidingNet [3], nazywany też Ikarus [4] jest systemem szacującym prawdopodobieństwo znalezienia komina termicznego. Na podstawie wielu zapisów lotów paralotni i lotni (rys 1.3) program analizuje częstotliwość występowania kominów termicznych na różnych obszarach, a następnie generuje mapę (rys 1.4). System ten jest nadal aktywnie rozwijany (stan na 2014 r.), i obecnie wykorzystuje ponad 750.000 zapisów lotów, obejmujących ponad 5.100.000 wykrytych kominów termicznych. Wygenerowane mapy można przeglądać w Internecie ¹ jako dodatkowe warstwy do map Google lub OpenStreetMap.



Rysunek 1.3: Drogi powietrzne zapisów z rejestratorów, ParaglidingNet



Rysunek 1.4: Mapa kominów termicznych, ParaglidingNet

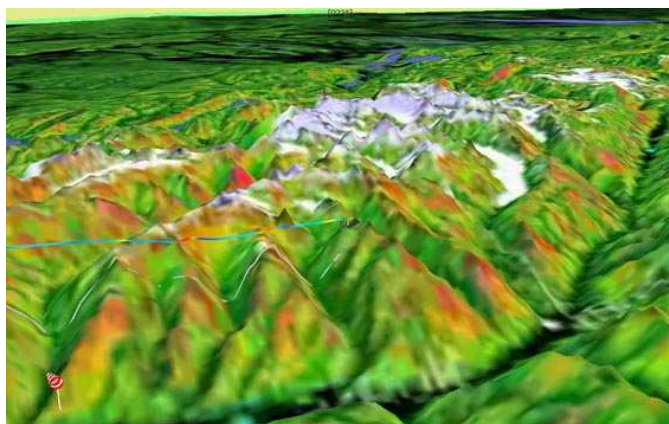
1.3.2. TherMap

TherMap ² (rys. 1.5) jest projektem wyznaczającym korzystne do szybowania obszary na podstawie topografii terenu, co istotne nie wykorzystując zapisów lotów. Ze strony projektu można pobrać mapy dla popularnych pasm górskich, w formacie grafiki rastrowej lub plików .kmz do Google Earth. Mapy są posegregowane ze względu na porę roku oraz porę dnia, jako że parametry te istotnie wpływają na wynik modelu.

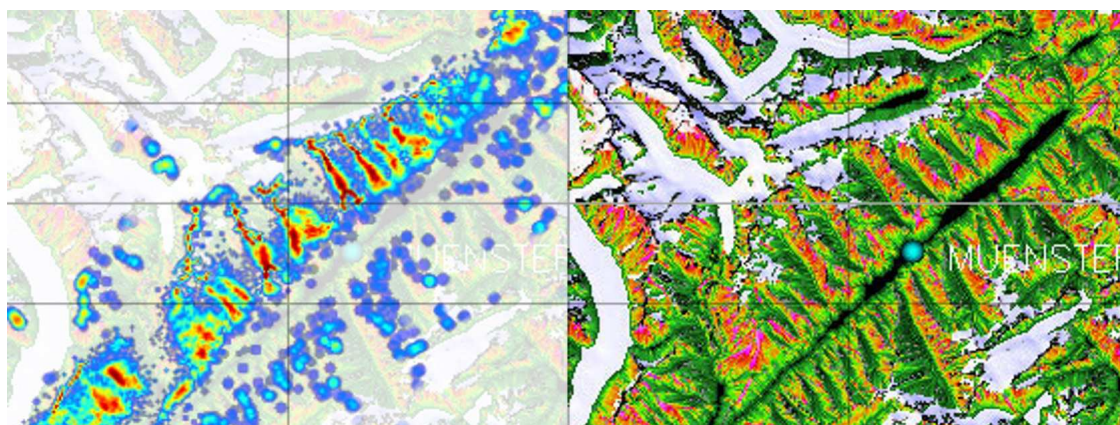
¹<http://thermal.kk7.ch/>

²<http://www.aerodrome-gruyere.ch/thermap/>

Wyniki tego modelu są zgodne z wynikami ParaglidingNet/Ikarus (rys 1.6), szczegóły można znaleźć w [3].



Rysunek 1.5: Mapy TherMap

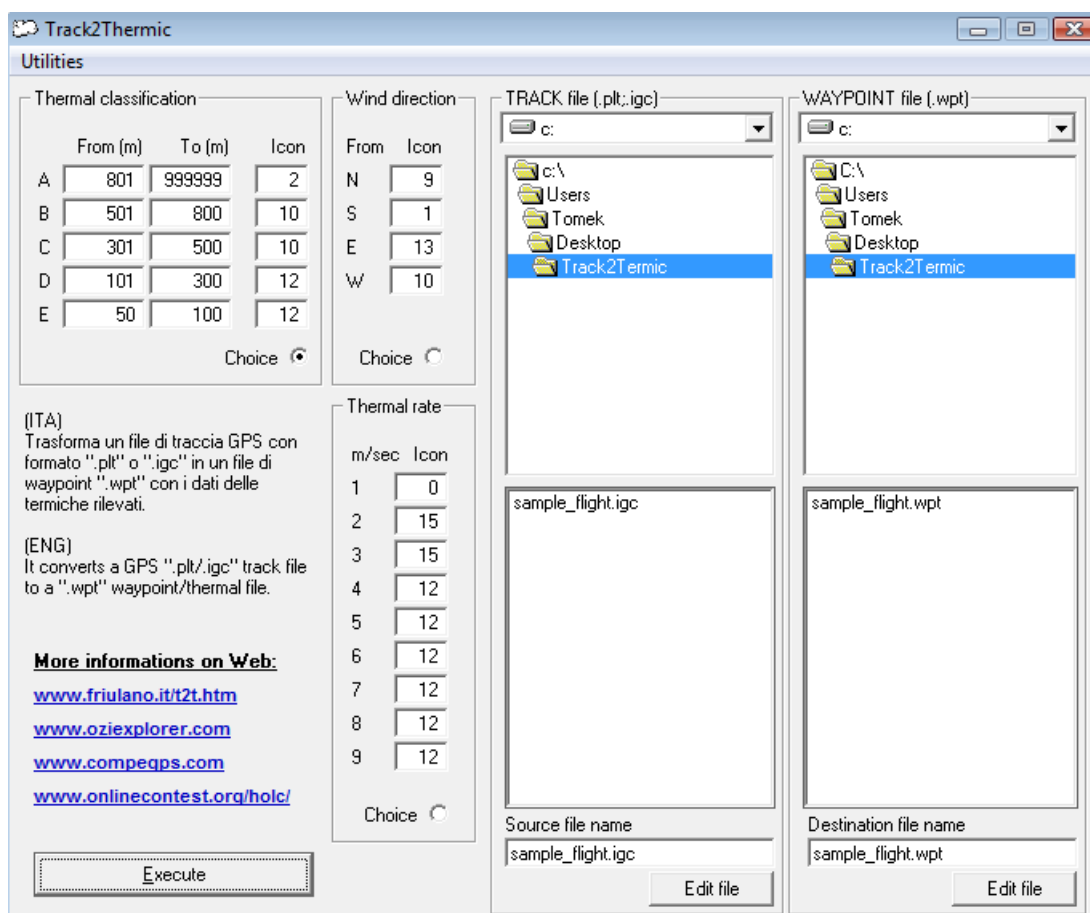


Rysunek 1.6: Porównanie ParaglidingNet z TherMap

1.3.3. Track2Thermic

Track2Thermic ³ (rys. 1.7) jest programem przeznaczonym głównie dla użytkowników paralotni. Program ten wykrywa kominy termiczne w zapisach lotów. Kominy te, po wstępnej segregacji na podstawie swoich podstawowych własności (prędkość pionowa, wysokość) mogą być wyeksportowane do pliku, w celu dalszego użycia na przenośnym urządzeniu nawigacyjnym podczas lotu. Ponadto autorzy programu utrzymują na swojej stronie internetowej bazę kominów termicznych występujących na terenie Europy. Baza ta jest wynikiem zastosowania programu na zapisach lotów ze strony OLC z lat 2002–2004. Najnowsza wersja programu pochodzi z listopada 2004 roku.

³<http://www.xweek.it/public/track2thermic/t2t.htm>



Rysunek 1.7: Program Track2Thermic

1.3.4. GlidingHotspots.eu

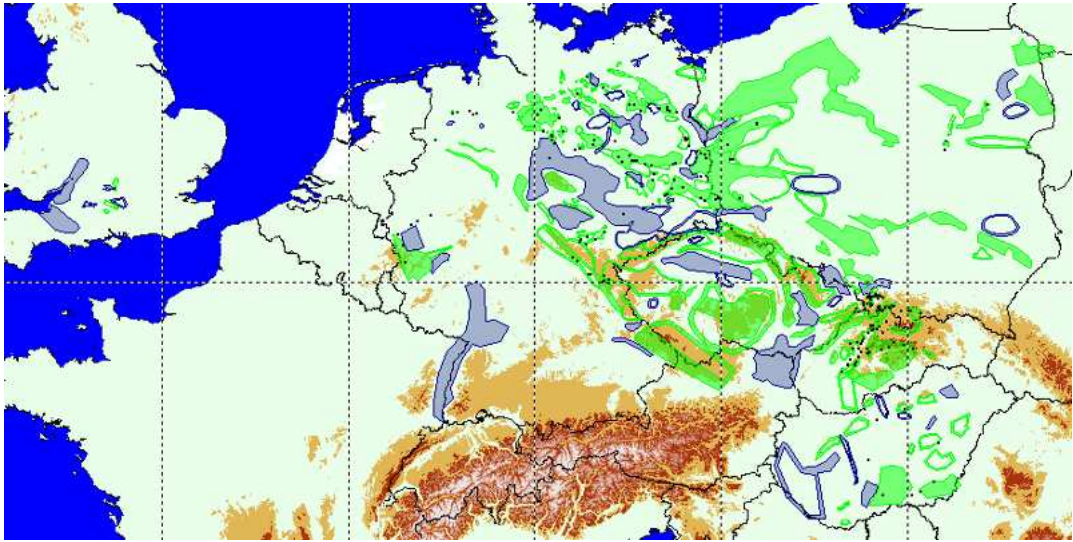
GlidingHotspots.eu⁴ jest ręcznie zarządzaną bazą danych ognisk termicznych (ang. *hotspots*) oraz obszarów o zwiększonej intensywności termicznej. Autorzy zachęcają użytkowników do przesyłania informacji na temat aktywności termicznej poprzez email. Gotowe mapy (rys. 1.8) są dostępne do pobrania na stronie projektu w formacie przystosowanym do wykorzystania w programie SeeYou.

1.3.5. Mapy Janusza Centki

Ośrodek szybowcowy w Lesznie pod Poznaniem jest jednym z wiodących ośrodków szybowcowych w Polsce. Okolice tego ośrodka zostały szczegółowo opisane na mapie (rys. 1.9) pod kątem przydatności do wykonywania lotów szybowcowych przez pilota szybowcowego Janusza Centkę, wielokrotnego zdobywcę tytułu Mistrza Świata oraz tytułu Mistrza Europy, pilota, który spędził w szybowcu tysiące godzin, z czego sporą część w samym tylko rejonie Leszna.

Mapę tę przetworzyłem, wydobywając jedynie lokalizacje istotnych lotnisk oraz obszary oznaczone jako dobre/złe. Wynikiem tej operacji jest rysunek 1.10.

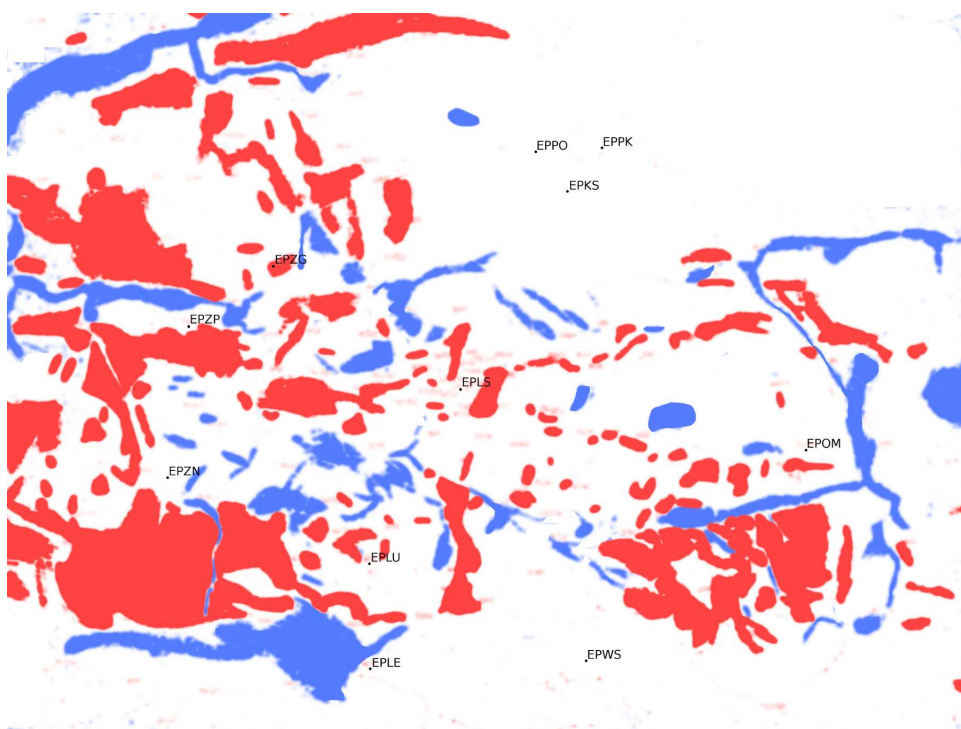
⁴<http://glidinghotspots.eu/>



Rysunek 1.8: GlidingHotspots.eu, mapa



Rysunek 1.9: Mapa Janusza Centki



Rysunek 1.10: Przetworzona mapa Janusza Centki

Rozdział 2

Mapy

2.1. Definicja mapy

Celem tej pracy jest stworzenie mapy wznoszeń termicznych, czyli mapy obrazującej wpływ terenu na jakość kominów termicznych. W kontekście mapy kominem termicznym będę nazywał trójkę

$$T_i = \langle lat_i, lon_i, vv_i \rangle \quad (2.1)$$

gdzie $lat_i \in [-90^\circ, 90^\circ]$ jest szerokością geograficzną, $lon_i \in [-180^\circ, 180^\circ]$ jest długością geograficzną, a $vv_i \in \mathbb{R}$ jest prędkością pionową komina (jakością), wyrażoną w metrach na sekundę (m/s).

Mapą będę nazywał dowolną funkcję postaci

$$M: D \rightarrow \mathbb{R} \quad (2.2)$$

gdzie $D = [-90^\circ, 90^\circ] \times [-180^\circ, 180^\circ]$ jest zbiorem dopuszczalnych współrzędnych geograficznych. Z punktu widzenia Machine Learning'u jest to predykcja prędkości pionowej komina przy zadanych współrzędnych geograficznych.

Błędem mapy na zbiorze kominów termicznych $\mathcal{T} = \{T_1, T_2, \dots, T_n\}$ będę nazywał wielkość

$$\text{MSE}(M, \mathcal{T}) = \frac{\sum_{i=1}^n (M(lat_i, lon_i) - vv_i)^2}{n} \quad (2.3)$$

Względnym błędem mapy na zbiorze kominów termicznych $\mathcal{T}$ będę nazywał wielkość

$$\text{MSE}_{\text{rel}}(M, \mathcal{T}) = \frac{\text{MSE}(M, \mathcal{T})}{\sigma_{\mathcal{T}}^2} \quad (2.4)$$

gdzie

$$\sigma_{\mathcal{T}}^2 = \frac{1}{n} \sum_{i=1}^n (\overline{vv_{\mathcal{T}}} - vv_i)^2 \quad (2.5)$$

$$\overline{vv_{\mathcal{T}}} = \frac{1}{n} \sum_{i=1}^n vv_i \quad (2.6)$$

2.2. Wielomiany interpolacyjne Hermite’a

Stawiając sobie za cel stworzenie mapy minimalizującej MSE_{rel} , ale jednocześnie dającej się w sensowny sposób znaleźć algorytmicznie, postanowiłem zawęzić zbiór map do klasy funkcji $C^r(D)$ kawałkami wielomianowych.

Wygodnym sposobem zapisu wielomianów jest zapis w *bazie Hermite’a* [8]. Bazy Hermite’a to ciąg $(B_1, B_3, B_5, \dots)$ baz wielomianów odpowiednio stopnia ≤ 1 (B_1), stopnia ≤ 3 (B_3), stopnia ≤ 5 (B_5), itd.; w których współczynniki wielomianu $w(x)$ to wartości oraz pochodne $w(x)$ w punktach 0 i 1. Na przykład współczynniki wielomianu $w(x)$ stopnia ≤ 3 w bazie B_3 to: $w^{(0)}(0)$, $w^{(0)}(1)$, $w^{(1)}(0)$, $w^{(1)}(1)$. Bazy te pozwalają na łatwe budowanie wielomianów o zadanych wartościach i pochodnych w dwóch punktach (0 i 1).

Bazowe wielomiany Hermite’a dla wielomianów stopnia ≤ 3 , czyli zbiór B_3 , to:

$$h_{00}(x) = 2x^3 - 3x^2 + 1 \quad (2.7)$$

$$h_{01}(x) = -2x^3 + 3x^2 \quad (2.8)$$

$$h_{10}(x) = x^3 - 2x^2 + x \quad (2.9)$$

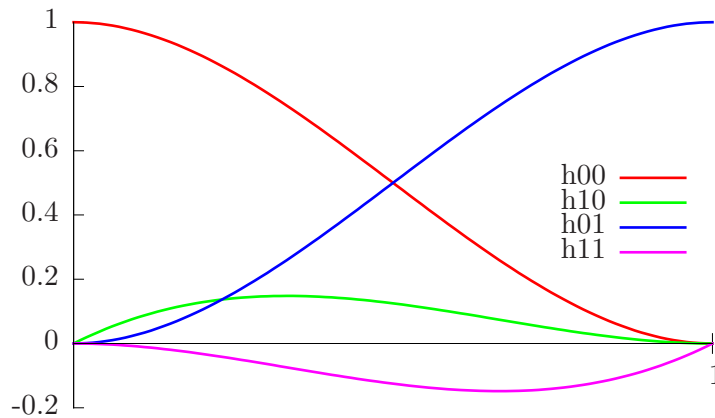
$$h_{11}(x) = x^3 - x^2 \quad (2.10)$$

Wówczas dla dowolnego wielomianu $w(x)$ stopnia ≤ 3 mamy:

$$\begin{aligned} w(x) = & w(0) \cdot h_{00}(x) \\ & + w(1) \cdot h_{01}(x) \\ & + w'(0) \cdot h_{10}(x) \\ & + w'(1) \cdot h_{11}(x) \end{aligned} \quad (2.11)$$

co można też zapisać w notacji macierzowej:

$$w(x) = \begin{bmatrix} h_{00}(x) \\ h_{01}(x) \\ h_{10}(x) \\ h_{11}(x) \end{bmatrix}^T \begin{bmatrix} w(0) \\ w(1) \\ \frac{\partial}{\partial x} w(0) \\ \frac{\partial}{\partial x} w(1) \end{bmatrix} \quad (2.12)$$



Rysunek 2.1: Bazowe wielomiany Hermite’a dla wielomianów stopnia ≤ 3

Współczynniki przy wielomianach bazowych h_{ij} można interpolować kolejnymi wielomianami, co prowadzi do uzyskania wielomianów wielu zmiennych. Wówczas mamy 2^k wierzchołków postaci $\{0, 1\}^k$, gdzie k to liczba zmiennych, i w każdym wierzchołku ustalamy

wartość wielomianu, oraz odpowiednie pochodne cząstkowe. Przykładowy opis wielomianu dwóch zmiennych w notacji macierzowej:

$$w(x, y) = \begin{bmatrix} h_{00}(x) \\ h_{01}(x) \\ h_{10}(x) \\ h_{11}(x) \end{bmatrix}^T \begin{bmatrix} w(0,0) & w(0,1) & \frac{\partial w}{\partial y}(0,0) & \frac{\partial w}{\partial y}(0,1) \\ w(1,0) & w(1,1) & \frac{\partial w}{\partial y}(1,0) & \frac{\partial w}{\partial y}(1,1) \\ \frac{\partial w}{\partial x}(0,0) & \frac{\partial w}{\partial x}(0,1) & \frac{\partial^2 w}{\partial x \partial y}(0,0) & \frac{\partial^2 w}{\partial x \partial y}(0,1) \\ \frac{\partial w}{\partial x}(1,0) & \frac{\partial w}{\partial x}(1,1) & \frac{\partial^2 w}{\partial x \partial y}(1,0) & \frac{\partial^2 w}{\partial x \partial y}(1,1) \end{bmatrix} \begin{bmatrix} h_{00}(y) \\ h_{01}(y) \\ h_{10}(y) \\ h_{11}(y) \end{bmatrix} \quad (2.13)$$

2.2.1. Zastosowanie do budowania mapy

Wyżej opisane wielomiany Hermite’a wielu zmiennych to elementy $C^r([0, 1]^k)$. Wielomiany te można skalować, otrzymując wielomiany opisane na innych kostkach. Ponadto wielomiany te można sklejać, otrzymując gładkie funkcje opisane na sumach kostek. Sklejając wielomiany należy użyć identycznych współczynników we wspólnych wierzchołkach kostek, oraz jednokowej stałej skalującej, zapewnia to gładkość otrzymanej funkcji na całej dziedzinie.

Założmy teraz, że zbiór $D = [-90^\circ, 90^\circ] \times [-180^\circ, 180^\circ]$ został podzielony na pewną liczbę przystających kostek (prostokątów). Wówczas zbiór F funkcji kawałkami wielomianowych otrzymanych poprzez sklejenie wielomianów Hermite’a jest podprzestrzenią liniową $C^r(D)$, a jej wymiar to $N = m \cdot (r + 1)^2$, gdzie m to ilość wierzchołków w podziale na kostki, a r to regularność. Niech $\mathcal{T} = \{T_1, T_2, \dots, T_n\}$ będzie zbiorem kominów termicznych. Będę szukał $f_{opt} \in F$ minimalizującego MSE_{rel} :

$$\begin{aligned} f_{opt} &= \underset{f \in F}{\operatorname{argmin}} MSE_{rel}(f, \mathcal{T}) \\ &= \underset{f \in F}{\operatorname{argmin}} MSE(f, \mathcal{T}) \\ &= \underset{f \in F}{\operatorname{argmin}} \frac{\sum_{i=1}^n (f(lat_i, lon_i) - vv_i)^2}{n} \\ &= \underset{f \in F}{\operatorname{argmin}} \sum_{i=1}^n (f(lat_i, lon_i) - vv_i)^2 \end{aligned} \quad (2.14)$$

Każda funkcja $f \in F$ ma zapis w bazie:

$$f = \sum_{j=1}^N x_j \cdot f_j \quad (2.15)$$

Szukanie f_{opt} jest równoważne szukaniu współczynników $\mathbf{x} \in \mathbb{R}^N$ optymalnej funkcji:

$$\underset{\mathbf{x} \in \mathbb{R}^N}{\operatorname{argmin}} \sum_{i=1}^n \left(\sum_{j=1}^N x_j \cdot f_j(lat_i, lon_i) - vv_i \right)^2 \quad (2.16)$$

Niech teraz $\mathbf{A} \in \mathbb{R}^{N,n}$ będzie macierzą, $\mathbf{x} \in \mathbb{R}^N$ będzie wektorem współczynników, a $\mathbf{b} \in \mathbb{R}^n$

będzie wektorem prędkości pionowych kominów termicznych:

$$\mathbf{A} = \begin{bmatrix} f_1(lat_1, lon_1) & f_2(lat_1, lon_1) & \dots & f_N(lat_1, lon_1) \\ f_1(lat_2, lon_2) & f_2(lat_2, lon_2) & \dots & f_N(lat_2, lon_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(lat_n, lon_n) & f_2(lat_n, lon_n) & \dots & f_N(lat_n, lon_n) \end{bmatrix} \quad (2.17)$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad (2.18)$$

$$\mathbf{b} = \begin{bmatrix} vv_1 \\ vv_2 \\ \vdots \\ vv_n \end{bmatrix} \quad (2.19)$$

Cel można teraz zwięźle zapisać jako:

$$\operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{Ax} - \mathbf{b}\|_2^2 \quad (2.20)$$

czyli instancję liniowego zadania najmniejszych kwadratów

2.2.2. Metoda LSQR

Dzięki zastosowaniu bazy Hermite’a macierz $\mathbf{A}$ będzie bardzo rzadka – w każdym wierszu będzie miała co najwyżej $(2r+2)^2$ niezerowych elementów. Do znalezienia $\mathbf{x} \in \mathbb{R}^N$ opisującego optymalną funkcję f_{opt} wykorzystałem standardową metodę numeryczną LSQR [6] (Sparse Equations and Least Squares), poprzez interfejs biblioteki SciPy [7] do obliczeń numerycznych w języku Python.

2.2.3. Przeuczanie

Zaproponowany algorytm minimalizuje błąd na danych treningowych, co może prowadzić do przeuczania. Aby zapobiec temu niekorzystnemu zjawisku zastosowałem dwa rodzaje „kar”:

1. odległość od średniej

$$\int_D (M(x, y) - \overline{vv})^2 d\mu(x, y) \quad (2.21)$$

2. oraz normę gradientu (pochodnej)

$$\int_D \left(\frac{\partial M}{\partial x}(x, y) \right)^2 + \left(\frac{\partial M}{\partial y}(x, y) \right)^2 d\mu(x, y) \quad (2.22)$$

Kary te można zapisać w sposób przybliżony jako dodatkowe wiersze instancji liniowego zadania najmniejszych kwadratów. I tak wyrażenie 2.21 w połączeniu z 2.16 można w sposób przybliżony zapisać jako:

$$\operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^N} \sum_{i=1}^n \left(\sum_{j=1}^N x_j \cdot f_j(\operatorname{lat}_i, \operatorname{lon}_i) - vv_i \right)^2 + \sum_l \left(\sum_{i=0}^N x_j \cdot \lambda \cdot f_i(\operatorname{lat}_l, \operatorname{lon}_l) - \lambda \cdot \overline{vv} \right)^2 \quad (2.23)$$

gdzie $\operatorname{lat}_l, \operatorname{lon}_l$ to punkty rozmieszczone jednorodnie na D , a $\lambda \geq 0$ jest parametrem kontrolującym wpływ kary. Natomiast wyrażenie 2.22 w połączeniu z 2.16 można w sposób przybliżony zapisać jako:

$$\begin{aligned} \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^N} \sum_{i=1}^n \left(\sum_{j=1}^N x_j \cdot f_j(\operatorname{lat}_i, \operatorname{lon}_i) - vv_i \right)^2 + \sum_l \left(\sum_{i=0}^N x_j \cdot \lambda \cdot \frac{\partial}{\partial x} f_i(\operatorname{lat}_l, \operatorname{lon}_l) - 0 \right)^2 \\ + \sum_l \left(\sum_{i=0}^N x_j \cdot \lambda \cdot \frac{\partial}{\partial y} f_i(\operatorname{lat}_l, \operatorname{lon}_l) - 0 \right)^2 \end{aligned} \quad (2.24)$$

gdzie $\operatorname{lat}_l, \operatorname{lon}_l, \lambda$ – podobnie jak poprzednio.

Kara typu „odległość od średniej” działa bardzo dobrze w praktyce, natomiast kara „norma gradientu” powoduje znaczne pogorszenie uwarunkowania macierzy A , i z tego powodu nie nadaje się do użycia, przynajmniej przy użyciu algorytmu LSQR.

2.3. k-NN

Metoda k-NN (ang. *k-Nearest Neighbors*) jest popularną metodą Machine Learning’u wykorzystywaną do klasyfikacji zbiorów danych opisanych parametrami ciągłymi [9]. Metoda ta wykorzystywana jest w dwóch podstawowych wariantach, to jest regresji (wynikiem jest liczba rzeczywista) oraz klasyfikacji (wynikiem jest etykieta z pewnego skończonego zbioru).

Zastosowanie tej metody wymaga metryki zdefiniowanej na zbiorze danych. W przypadku map termiki szybowcowej wykorzystałem metrykę odległości na sferze (*great-circle distance*), ponieważ danymi są tu współrzędne geograficzne:

$$d((\phi_1, \lambda_1), (\phi_2, \lambda_2)) = \arccos(\sin(\phi_1) \sin(\phi_2) + \cos(\phi_1) \cos(\phi_2) \cos(\lambda_2 - \lambda_1)). \quad (2.25)$$

Powyższy wzór ma równoważną postać o lepszych własnościach numerycznych dla małych wartości odległości [13]:

$$d((\phi_1, \lambda_1), (\phi_2, \lambda_2)) = 2 \arcsin \left(\sqrt{\sin^2 \left(\frac{\phi_2 - \phi_1}{2} \right) + \cos \phi_1 \cos \phi_2 \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right)} \right) \quad (2.26)$$

stąd też w obliczeniach stosowałem wzór 2.26.

Metoda k-NN dla parametru $k > 1$ potrzebuje sposobu uśrednienia wartości sąsiadów. Standardowe sposoby uśredniania dla metody k-NN w wariancie regresji to obliczanie średniej arytmetycznej z wartości sąsiadów lub średniej ważonej, dla wag dobieranych na podstawie odległości. Po przeprowadzeniu eksperymentów na pozyskanych danych wybrałem funkcję przypisującą wagi sąsiadom na podstawie ich odległości w kilometrach:

$$w(d) = \frac{1}{d + 1} \quad (2.27)$$

Wówczas zbiorowi sąsiadów $\{(vv_1, d_1), (vv_2, d_2), \dots, (vv_k, d_k)\}$ przypisujemy predykcję:

$$vv = \frac{\sum_{i=1}^k vv_i \cdot w(d_i)}{\sum_{i=1}^k w(d_i)} \quad (2.28)$$

Obliczanie predykcji metodą k-NN jest dosyć kosztowne obliczeniowo – w naiwnej implementacji algorytmu każde zapytanie ma złożoność $O(N)$, gdzie N to ilość krotek danych użytych do uczenia. Dzięki wykorzystaniu właściwości metryki, a w szczególności nierówności trójkąta, możliwe jest obniżenie tej złożoności kosztem jednorazowej konstrukcji specjalnych drzew. W implementacji użyłem algorytmów z bibliotek *mlpack* [10] oraz *scikit-learn* [11], które dobrze radziły sobie z wykorzystywanymi danymi.

Rozdział 3

Dane

3.1. Rejestratory, format IGC

Piloci szybowcowi masowo wykorzystują rejestratory GPS (ang. *GPS logger*), na przykład podczas wszelkiego rodzaju zawodów szybowcowych, czy podczas prób zdobycia odznak szybowcowych lub prób pobicia rekordów. Urządzenia te w ustalonych interwałach czasowych (1 do 10 sekund) zapisują krotki danych opisujące aktualny stan statku powietrznego. Pojedyncza krotka danych jest nazywana *fixem*.

Celem zapewnienia standardu w wymianie informacji z rejestratorów, a także sprecyzowania wymogów, jakie muszą spełniać rejestratory GPS, międzynarodowa organizacja lotnictwa sportowego *Fédération Aéronautique Internationale* (FAI, ang. *The World Air Sports Federation*), a konkretniej grupa *International Gliding Commission* (IGC) stworzyła format plików IGC służący do przechowywania zapisów lotów z certyfikowanych rejestratorów lotu [5].

Format IGC jest tekstowym formatem danych, w którym zapisy składają się z ciągu linii (rekordów). Każda linia rozpoczyna się wielką literą alfabetu łacińskiego opisującą typ rekordu, i kończy się znakami CR-LF. Na początku pliku IGC znajduje się linia opisująca producenta, model i numer seryjny rejestratora (rekord A). Dalej mamy linie nagłówkowe (rekordy H), zawierające informacje takie jak:

- data lotu
- imię i nazwisko pilota
- model szybowca
- numer rejestracyjny szybowca
- szczegółowe dane na temat użytych sensorów (GPS, mierniki ciśnienia)
- zadeklarowaną trasę lotu (rekord C)
- użyte rozszerzenia (rekordy I, J)

Po nagłówku następuje ciąg *fixów* (rekordy B), przeplatanych z dodatkowymi danymi (rekordy E, K, F). Każdy z *fixów* zawiera informacje takie jak:

- czas GPS
- współrzędne GPS
- wysokość GPS nad poziom morza

- wysokość ciśnieniową (z czujnika ciśnieniowego)
- różne inne dane, w zależności od modelu rejestratora i włączonych rozszerzeń

Na końcu pliku IGC znajduje się rekord G – podpis cyfrowy, składany przez rejestrator po zakończeniu lotu.

3.2. Zestawienie źródeł danych

Zawody szybowcowe dzielą się na dwie duże grupy: zawody tradycyjne oraz zawody korespondencyjne.

Zawody tradycyjne odbywają się w ustalonym miejscu i trwają od jednego do trzech tygodni, podczas których organizatorzy codziennie analizują prognozę pogody, ustalają trasy dnia, a następnie piloci ścigają się na uzgodnionych trasach. Po zakończonych lotach piloci zdają sędziom cyfrowo podpisane pliki z rejestratorów, i na tej podstawie otrzymują punkty. Przykładami takich zawodów są mistrzostwa krajowe, mistrzostwa Europy, mistrzostwa Świata, lokalne zawody szybowcowe.

Zawody korespondencyjne typowo trwają cały rok (cały sezon). Ustalanie tras jest pozostawione do decyzji pilota, a punkty są przyznawane za przelecianą odległość, czasami także za uzyskaną prędkość, na podstawie zapisów z rejestratorów przesyłanych przez Internet. Przykładami zawodów korespondencyjnych są *On-line Contest (OLC)*, czy *Szybowcowy Puchar Polski (SzPP)*.

Zapisy lotów wraz z ocenami sędziów są publikowane na stronach internetowych zawodów. Dzięki temu wszyscy zainteresowani mogą analizować loty zawodników, co jest bardzo przydatne do porównawczej analizy strategii użytych podczas lotu. Zapisy lotów zawodów tradycyjnych są zazwyczaj publikowane w serwisie internetowym *SoaringSpot*, zaś zapisy z zawodów korespondencyjnych na stronach internetowych tych zawodów. Dostępne zapisy z lat 2003-2013 podsumowuje tabela 3.1.

Tabela 3.1: Zestawienie publicznych serwisów z plikami IGC, lata 2003 - 2013

Serwis	Ilość zapisów	Link
SoaringSpot	137 000	http://www.soaringspot.com/
On-line Contest	681 000	http://www.onlinecontest.org/
Szybowcowy Puchar Polski	4 500	http://szpp.pl/

Wykorzystałem wszystkie zapisy z SzPP oraz SoaringSpot, a zatem łącznie około 140 000 zapisów lotów szybowców. W niniejszym opracowaniu nie użyłem danych z serwisu OLC, ponieważ zgodnie z warunkami użycia serwisu pojedynczy użytkownik może ściągnąć maksymalnie 10 zapisów dziennie, co jest ilością niewystarczającą do pobrania sensownej ilości zapisów.

3.3. Wykrywanie kominów termicznych

Dane z rejestratora są bardzo bogatym źródłem informacji. Jednakże do celów stworzenia mapy termiki postanowiłem wyodrębnić z tych danych jedynie informacje o krążeniu w kominach termicznych.

Aby oznaczyć fragmenty zapisu odpowiadające za krążenie wykorzystałem proste odcięcie prędkości kątowej szybowca. Założyłem, że pilot nie ma żadnych sensownych powodów do krążenia poza kominami termicznymi, a zatem wszystkie wykryte krążenia oznaczają komin

termiczny. Zastosowałem progową wartość prędkości kątowej $4^\circ/s$, a następnie tak otrzymany ciąg binarnych oznaczeń *krąży/nie krąży* przepuściłem przez łańcuch Markowa (algorytm Viterbiego). Łańcuch Markowa wygładził uzyskane oznaczenia, wówczas przyjąłem, że komin termiczny odpowiada maksymalnemu spójnemu podciągowi oznaczeń *krąży*. Ponadto skupiłem się wyłącznie na kominach w których pilot spędził co najmniej 60 sekund, co pozwoliło wyeliminować oznaczanie momentów zawracania czy powolnych zakrętów jako krążenia.

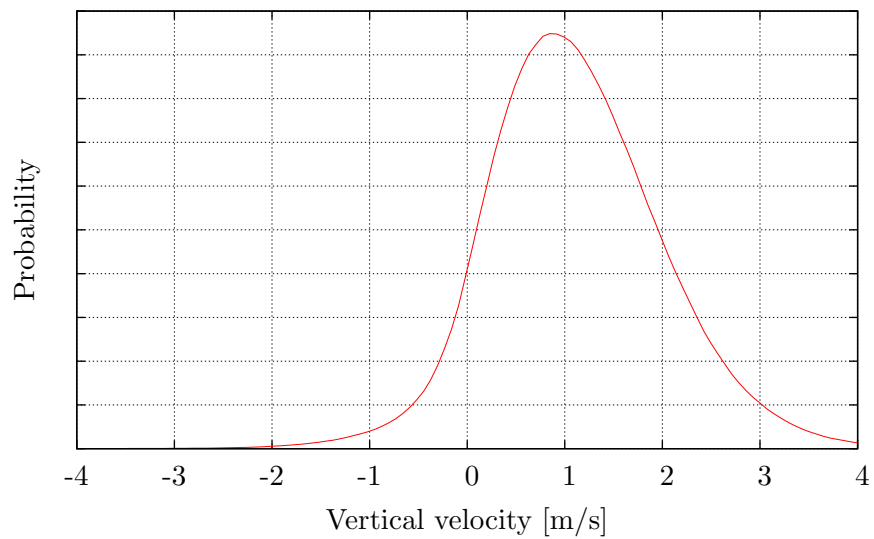
Dodatkowo konieczne było oznaczenie momentu startu i lądowania – ze względu na niedokładność odbiorników GPS szybowiec stojący w miejscu będzie na zapisie rejestratora się losowo przesuwiał w zasięgu kilkunastu metrów od swojej rzeczywistej pozycji. Zachowanie takie mogłoby zostać omyłkowo zinterpretowane jako krążenie, ponieważ mogłoby spełniać warunek zmiennej prędkości kątowej. W celu wykrycia momentów startu i lądowania zastosowałem metodę analogiczną do wykrywania krążenia – proste odcięcie prędkości liniowej (próg 40 km/h), połączone z wygładzaniem łańcuchem Markowa.

Podczas eksperymentów z wykorzystanymi metodami budowania map okazało się, że komin termiczny znajdujące się blisko miejsca startu szybowca mają na ogół niższe wartości prędkości pionowej niż pozostałe. Wyniki te nie były zgodne z intuicją, i są najprawdopodobniej skutkiem specyfiki lotów zawodniczych, w których ważnym elementem strategii pilota jest wybór momentu odejścia na trasę. W przypadkach gdy piloci spodziewają się poprawy pogody w późniejszej części dnia, i spodziewany czas potrzebny na przelecenie całej trasy jest krótszy od spodziewanego czasu występowania zjawisk termicznych, korzystnym może okazać się oczekiwanie w pobliżu miejsca startu. W takich przypadkach wyszukiwanie najsilniejszych wznoszeń nie jest priorytetem pilota, a nawet jest odrobinę niekorzystne – w końcu celem takiego zachowania jest po prostu pozostanie w powietrzu, w rejonie linii startowej. Aby wyeliminować wpływ tego zjawiska usunąłem ze zbioru wykrytych kominów termicznych wszystkie znajdujące się w odległości mniejszej niż 20 km od punktu startu. Tę minimalną odległość dobrałem intuicyjnie na podstawie wykresu średniej prędkości komina termicznego w zależności od odległości od punktu startu (rys. 3.4).

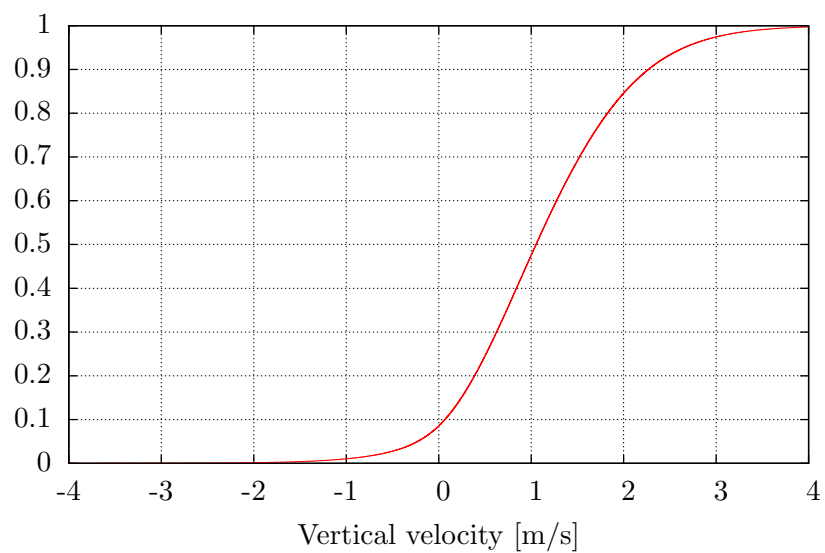
3.4. Ogólne statystyki danych

Przetworzyłem 140 000 zapisów lotów, o łącznej wielkości 32 GB. Zapisy obejmowały lata 2003-2013. Algorytm wykrywania krążenia oznaczył łącznie 3 500 000 kominów termicznych. Większość lotów odbywała się w Europie, chociaż jest też trochę zapisów z Australii i Nowej Zelandii, Ameryki Północnej (Teksas), Ameryki Południowej (Chile) i Azji (głównie Rosja).

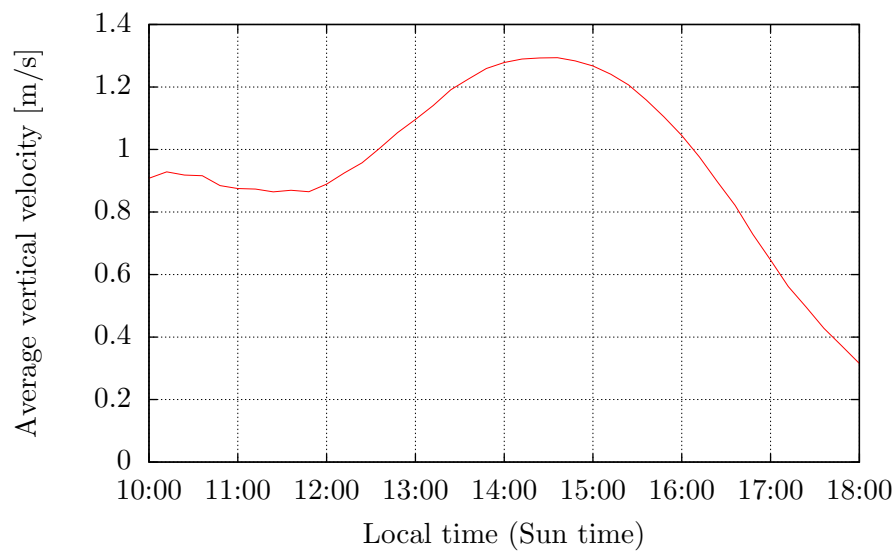
Na uwagę zasługuje tu rysunek 3.5. Pokazuje on, że czas krążenia w kominie termicznym ma w przybliżeniu rozkład wykładniczy. Istotną własnością rozkładu wykładniczego jest *brak pamięci*, to znaczy następująca własność: jeżeli X jest zmienną losową o rozkładzie wykładniczym, to $\forall_{t,s>0} \Pr(X > t) = \Pr(X > t + s \mid X > s)$. W przypadku rozkładu czasu krążenia oznacza to, że czas już spędzony w kominie statystycznie nie wpływa na decyzję pilota o opuszczeniu komina.



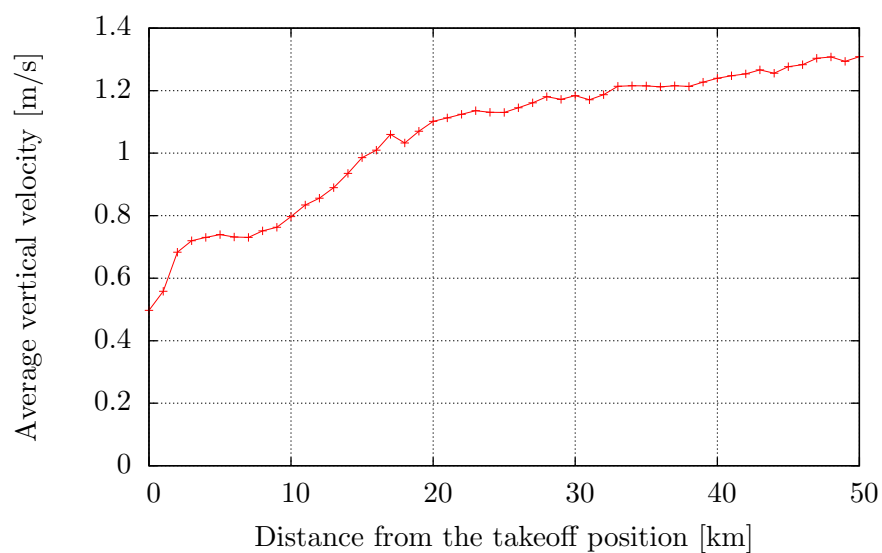
Rysunek 3.1: Estymator jądrowy gęstości – prędkość pionowa komina termicznego



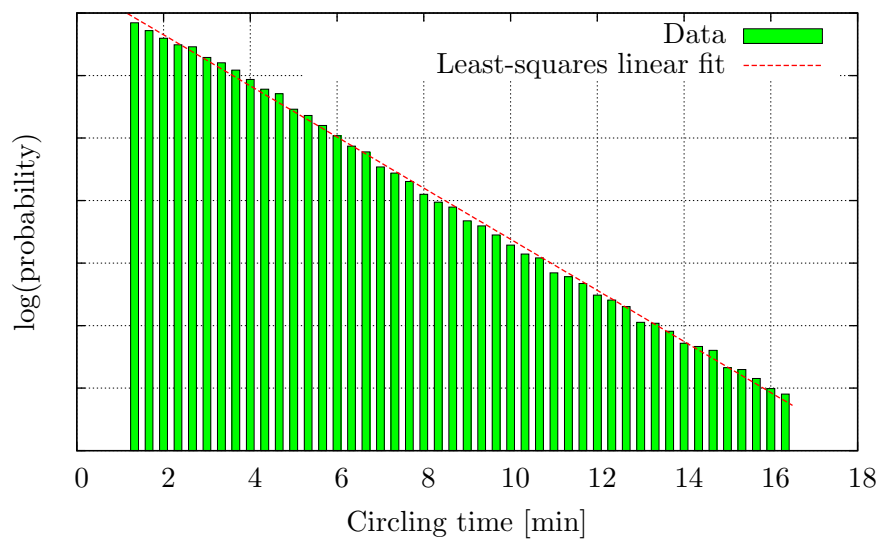
Rysunek 3.2: Dystrybuanta empiryczna – prędkość pionowa komina termicznego



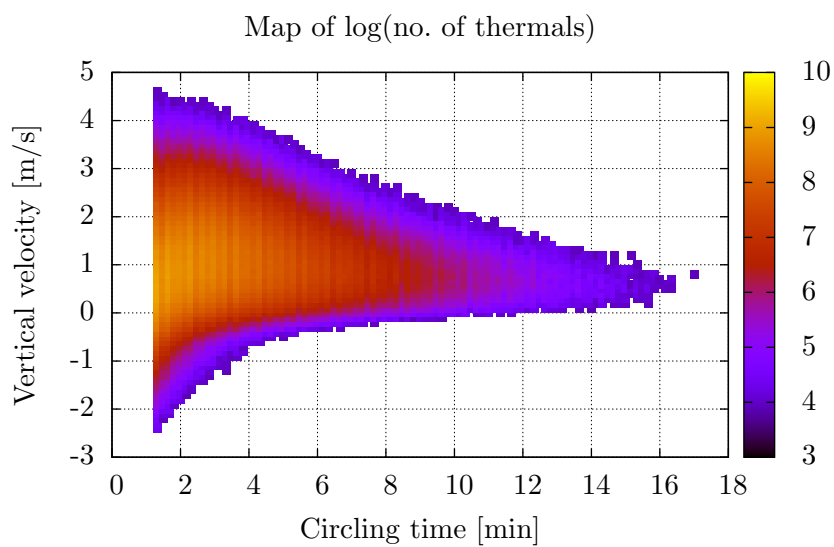
Rysunek 3.3: Średnia prędkość komina w ciągu dnia (czas lokalny, słoneczny)



Rysunek 3.4: Średnia prędkość komina termicznego w funkcji odległości od startu



Rysunek 3.5: Zlogarytmowany histogram czasu krążenia w kominie – rozkład wykładniczy



Rysunek 3.6: Zlogarytmowany histogram łącznego rozkładu prędkości pionowej i czasu krążenia w kominie termicznym

Rozdział 4

Wyniki eksperymentalne

4.1. Pomiar skuteczności klasyfikatorów

Eksperymenty z danymi przeprowadziłem z wykorzystaniem chmury obliczeniowej *Amazon EC2* oraz aplikacji do zarządzania zadaniami Jenkins. Do pomiaru jakości klasyfikatorów użyłem względnego błędu mapy, MSE_{rel} (równanie 2.4), wykorzystując zasadę walidacji krzyżowej (*cross-validation*). Dla każdej rundy zbiór wykrytych kominów termicznych dzieliłem na 5 podzbiorów, klasyfikatory budowałem w oparciu o wybrane 4, i testowałem na pozostałym podziorze. W ten sposób klasyfikator otrzymywał 5 wyników (jeden dla każdego podziału na 4 podzbiory treningowe i 1 testowy) na każdą rundę. Wyniki uzyskane wewnątrz rundy następnie uśredniałem, otrzymując częściowe średnie. Procedurę tę powtarzałem 5 razy, co dawało 25 wyników, pięć częściowych średnich, po jednej na rundę.

Podzbiory danych były tworzone w taki sposób, aby dane zarejestrowane konkretnego dnia były w całości zawarte w jednym z podzbiorów. Takie zachowanie było konieczne aby wyeliminować wpływ silnego podobieństwa danych pochodzących z „peletonu” szybowców (zjawiska dosyć częstego na zawodach szybowcowych) na wynik klasyfikatora.

Tabela 4.1 podsumowuje wyniki klasyfikatora opartego o wielomiany interpolacyjne Hermite’a, zaś tabela 4.2 podsumowuje wyniki klasyfikatora k-NN. Wyniki klasyfikatorów MSE_{rel} są ilorazami dwóch porównywalnych wielkości (równanie 2.4), więc dla wygody zostały podane w procentach.

Tabela 4.1: Wyniki klasyfikatora opartego na wielomianach interpolacyjnych Hermite’a

Runda CV	Średnia MSE_{rel}
1	93.29 %
2	93.64 %
3	93.72 %
4	93.89 %
5	93.88 %
Łączna średnia	93.68 %

Klasyfikator k-NN osiągnął trochę lepsze rezultaty niż klasyfikator oparty na wielomianach Hermite’a, kosztem jednak sporo większej złożoności obliczeniowej. Wykonanie obliczeń potrzebnych do zbudowania tabeli 4.1 trwało około 5 godzin na instancji Amazon EC2 *m1.medium*, natomiast wykonanie obliczeń potrzebnych do stworzenia tabeli 4.2 trwało ponad 42 godziny.

Tabela 4.2: Wyniki klasyfikatora k-NN z wagami dobieranymi na podstawie odległości

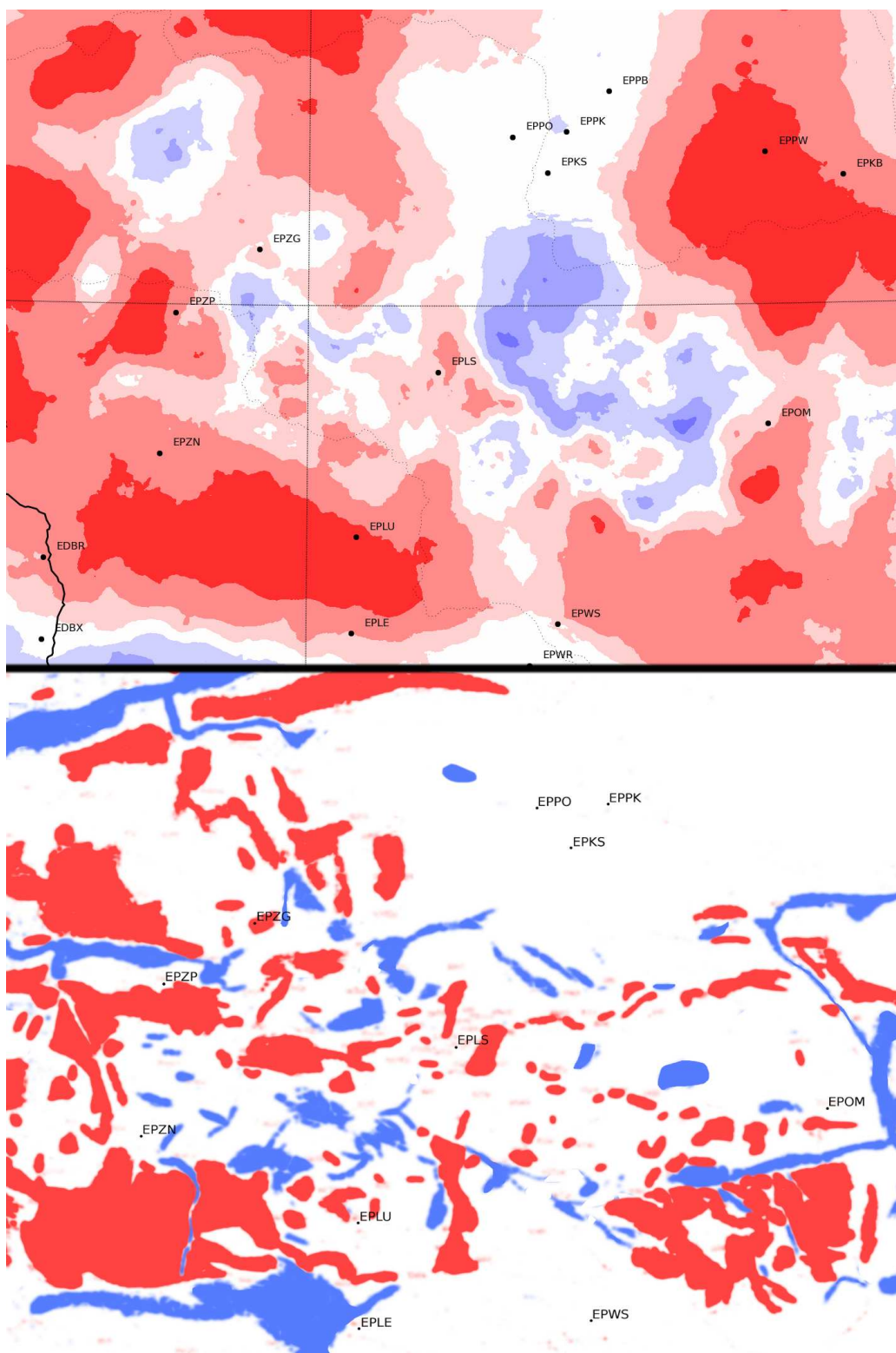
Runda CV	MSE_{rel}
1	92.40 %
2	92.70 %
3	92.88 %
4	92.95 %
5	93.12 %
Łączna średnia	92.81 %

4.2. Mapy termiki szybowcowej

Mapy tworzyłem budując klasyfikatory na kompletnym zbiorze wykrytych kominów termicznych, a następnie nanosząc odpowiedzi klasyfikatorów na obrazki. Do wizualizacji wyników użyłem biblioteki *matplotlib* [12], a konkretnie podmodułu *basemap*, służącego do wizualizacji danych geograficznych. Dla wygody użytkownika naniosłem na mapy położenia oraz kody ICAO znanych lotnisk.

Rysunek 4.1 przedstawia wyniki klasyfikatora k-NN, natomiast rysunek 4.2 przedstawia wyniki klasyfikatora opartego na wielomianach interpolacyjnych Hermite’a. Widać tu problemy klasyfikatora opartego na wielomianach Hermite’a: wyniki nie są jednorodne, widoczna jest tu siatka punktów kontrolnych ułożonych równoległe do równoleżników i południków. Klasyfikator k-NN nie przejawia tego typu problemów.

Powstałe mapy porównałem wizualnie z różnymi mapami wymienionymi w sekcji 1.3. Porównanie map modelu k-NN z mapami Janusza Centki (sekcja 1.3.5) dla regionu okolic Leszna (EPLS) jest na rysunku 4.3. Należy zaznaczyć, że mapy Janusza Centki ograniczają się do przydzielenia każdemu punktowi jednej z trzech kategorii: dobrze, neutralnie, źle; natomiast mapy modelu k-NN są w ogólności dokładniejsze, i mają więcej kategorii. Pomimo tej istotnej różnicy widać podobieństwo w lokalizacji i kształcie rejonów aktywnych termicznie.



Rysunek 4.3: Porównanie wyników klasyfikatora k-NN (u góry) z mapami Janusza Centki (u dołu)

Rozdział 5

Podsumowanie i perspektywy dalszych prac

Niniejsza praca potwierdza istnienie związku pomiędzy lokalizacją komina termicznego a jego prędkością pionową, dostarczając czytelnych map obrazujących wpływ lokalizacji na jakość komina termicznego. Mapy te są zgodne z mapami stworzonymi na podstawie opinii doświadczonych ekspertów, a modele matematyczne leżące u podstaw tych map przechodzą pozytywnie weryfikację ich własności w procesie walidacji krzyżowej, tłumacząc 7-8% wariancji prędkości pionowej kominów termicznych.

Dane pochodzące z zapisów lotów IGC są bogatym źródłem informacji. W niniejszej pracy wykorzystałem jedynie informacje o lokalizacji i prędkości pionowej komina termicznego. Innymi przydatnymi informacjami które można odczytać z zapisu lotu są:

- kierunek i siła wiatru – na przykład na podstawie dryfowania szybowca krążącego w kominie;
- data wykonywania lotu, co można połączyć z historycznymi informacjami meteorologicznymi;
- model szybowca;
- zaplanowana trasa.

Wykorzystanie tych dodatkowych informacji może poprawić wyniki modeli, oraz w rezultacie przyczynić się do jeszcze lepszego zrozumienia zagadnienia prądów termicznych.

Ponadto ciekawym jest zagadnienie stworzenia mapy termiki szybowcowej wyłącznie w oparciu o topografię terenu. Podejście takie prezentuje TherMap (sekcja 1.3.2), koncentrując się na terenach górskich. Dane z rejestratorów szybowcowych mogą okazać się bardzo przydatne w rozrozu tego typu modeli jako warstwa weryfikacji wyników modelu topograficznego.

Poza prądami termicznymi piloci szybowcowi wykorzystują górskie zjawiska żagla oraz fali. Zjawiska te są bardzo ściśle powiązane z działaniem wiatru i kształtem terenu, stąd też bardzo skutecznym może okazać się stworzenie analogicznych map dla tych właśnie zjawisk. Zjawiska te będąc istotnie różne od zjawiska termiki nie zostały poruszone w niniejszym opracowaniu.

Bibliografia

- [1] Władysław Parczewski
Meteorologia szybowcowa
Wydawnictwo Ligi Lotniczej, Warszawa 1953
- [2] Andrzej Abłamowicz
Podręcznik pilota szybowcowego, rozdział szósty.
Wydawnictwa Komunikacji i Łączności, Warszawa 1967
- [3] Michael von Kaenel
ParaglidingNet: A Sensor Network for Thermal Research
Master's Thesis, Swiss Federal Institute of Technology Zurich, August 2010
<http://thermal.kk7.ch/>
- [4] Michael von Kaenel, Philipp Sommer, Roger Wattenhofer
Ikarus: Large-Scale Participatory Sensing at High Altitudes
Proceedings of the 12th Workshop on Mobile Computing Systems and Applications: 63–68, ACM, 2011
- [5] Fédération Aéronautique Internationale, International Gliding Commission
Technical Specification for GNSS Flight Recorders
<http://www.fai.org/igc-documents>
- [6] Christopher Paige, Michael Saunders
LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares
ACM Transactions on Mathematical Software (TOMS), Volume 8 Issue 1, ACM, March 1982
<http://web.stanford.edu/group/SOL/software/lsqr/>
- [7] Eric Jones, Travis Oliphant, Pearu Peterson, et al.
SciPy: Open Source Scientific Tools for Python
<http://www.scipy.org/>
- [8] A. Spitzbart
A Generalization of Hermite's Interpolation Formula
American Mathematical Monthly 67 (1): 42–46, MAA, January 1960

- [9] N. S. Altman
An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression
 The American Statistician 46 (3): 175–185, ASA, August 1992
- [10] Ryan R. Curtin, James R. Cline, Neil P. Slagle, William B. March, P. Ram, Nishant A. Mehta, Alexander G. Gray
MLPACK: A Scalable C++ Machine Learning Library
<http://mlpack.org/>
 Journal of Machine Learning Research 14: 801–805, JMLR, 2013
- [11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay
Scikit-learn: Machine Learning in Python
 Journal of Machine Learning Research 12: 2825–2830 JMLR, 2011
<http://scikit-learn.org/>
- [12] John D. Hunter
Matplotlib: A 2D graphics environment
 Computing In Science & Engineering, vol. 9, no. 9, p. 90–95, IEEE, 2007
<http://matplotlib.org/>
- [13] R. W. Sinnott
Virtues of the Haversine
 Sky and Telescope, vol. 68, no. 2, 1984, p. 159