

Prototype d'Algorithme de Correction de l'Allure en Course à Pied

Achraf Lassoued

21 avril 2025

Table des matières

1	Objectif du Travail	2
2	Revue Bibliographique Approfondie	2
2.1	Strava's Grade Adjusted Pace (GAP)	2
2.2	Formule de Course de Daniels et Tables VDOT	2
2.3	Modèle de Coût Énergétique de Minetti (2002)	2
2.4	Modèles Basés sur l'Apprentissage Automatique	3
3	Analyse Comparative des Approches	3
4	Choix des Approches et Justification	3
4.1	Modèle de Coût Énergétique de Minetti	3
4.2	Modèles d'Apprentissage Automatique	3
5	Processus du Prototype	4
5.1	Étape 1 : Extraction des Données du Fichier GPX	4
5.2	Étape 2 : Application du Modèle de Minetti	4
5.3	Étape 3 : Intégration des Facteurs Environnementaux via Apprentissage Automatique	4
5.4	Étape 4 : Modélisation par Apprentissage Automatique	4
5.5	Étape 5 : Calcul des Temps Totaux	5
5.6	Étape 6 : Comparaison des Deux Approches	5
6	Analyse et Comparaison des Résultats	5
6.1	Résumé des Temps Obtenus	5
6.2	Écart entre les approches	5
6.3	Interprétation et Choix Méthodologique	6

1 Objectif du Travail

L'objectif principal est de concevoir un prototype fonctionnel d'un algorithme capable de corriger l'allure ou le temps de course en intégrant les variations de terrain (dénivelé) et les conditions environnementales (température, humidité). Ce prototype servira de base pour le développement futur d'un modèle plus avancé destiné à une utilisation en production.

2 Revue Bibliographique Approfondie

2.1 Strava's Grade Adjusted Pace (GAP)

Strava propose une fonctionnalité appelée GAP (Grade Adjusted Pace) qui ajuste l'allure d'un coureur en fonction de la pente du terrain. Bien que l'algorithme exact soit propriétaire, des analyses ont été réalisées pour comprendre son fonctionnement. Il semble que Strava utilise un facteur d'ajustement basé sur la pente, avec des ajustements plus importants pour les pentes plus raides.

Avantages :

- Intégration directe dans une plateforme populaire.
- Fournit une estimation rapide de l'allure corrigée.

Inconvénients :

- Manque de transparence sur l'algorithme utilisé.
- Ne prend pas en compte les facteurs environnementaux autres que la pente.

2.2 Formule de Course de Daniels et Tables VDOT

Développée par Jack Daniels, cette méthode utilise des tables VDOT pour estimer les performances d'un coureur en fonction de sa VO_2 max. Elle permet de déterminer des allures d'entraînement personnalisées et de prédire les performances sur différentes distances.

Avantages :

- Basé sur des données physiologiques solides.
- Permet une personnalisation des entraînements.

Inconvénients :

- Ne prend pas en compte les variations de terrain ou les conditions environnementales.
- Suppose une économie de course constante, ce qui peut varier entre les individus.

2.3 Modèle de Coût Énergétique de Minetti (2002)

Minetti et al. ont étudié le coût énergétique de la course en fonction de la pente. Leur modèle montre une relation non linéaire entre la pente et le coût énergétique, avec des coûts plus élevés pour les pentes raides, qu'elles soient ascendantes ou descendantes.

Avantages :

- Fournit une base physiologique pour ajuster l'allure en fonction de la pente.
- Applicable à la fois pour les montées et les descentes.

Inconvénients :

- Nécessite des données d'altitude précises.
- Ne prend pas en compte d'autres facteurs environnementaux comme la température ou le vent.

2.4 Modèles Basés sur l'Apprentissage Automatique

Des approches récentes utilisent l'apprentissage automatique pour prédire les performances de course en fonction de divers facteurs environnementaux. Ces modèles peuvent intégrer des variables telles que la température, l'humidité, la vitesse du vent et le terrain pour estimer la dépense énergétique et les stratégies d'allure optimales.

Avantages :

- Capables de gérer des relations complexes et non linéaires entre les variables.
- Adaptables aux données individuelles des coureurs pour des prédictions personnalisées.

Inconvénients :

- Nécessitent de grands ensembles de données pour l'entraînement et la validation.
- Peuvent manquer de transparence, rendant difficile l'interprétation de l'influence des variables individuelles.

3 Analyse Comparative des Approches

TABLE 1 – Comparaison des différentes approches

Approche	Ajustement	Facteurs Env.	Personnalisation	Transparence
Strava's GAP	Oui	Non	Non	Non
Formule de Daniels (VDOT)	Non	Non	Oui	Oui
Modèle de Minetti	Oui	Non	Non	Oui
Apprentissage Automatique	Oui	Oui	Oui	Non

4 Choix des Approches et Justification

Pour le développement du prototype, deux approches ont été retenues :

4.1 Modèle de Coût Énergétique de Minetti

Justification :

- **Simplicité** : Facile à implémenter avec des données de pente.
- **Précision** : Basé sur des études physiologiques solides.
- **Pertinence** : Spécifiquement conçu pour ajuster l'allure en fonction du dénivelé.
- **Transparence** : Formule claire et bien documentée.

4.2 Modèles d'Apprentissage Automatique

Justification :

- **Adaptabilité** : Capables de gérer des relations complexes entre multiples variables.
- **Personnalisation** : Permettent des prédictions adaptées aux individus.
- **Innovation** : Utilisent des techniques modernes prometteuses dans le domaine de la performance sportive.

5 Processus du Prototype

5.1 Étape 1 : Extraction des Données du Fichier GPX

Entrée : Fichier GPX contenant les données de parcours (latitude, longitude, altitude, timestamp).

Traitement :

- Lecture du fichier GPX à l'aide de la bibliothèque `gpxpy`.
- Extraction des points GPS et calcul des distances entre points successifs.
- Calcul des différences d'altitude pour déterminer la pente de chaque segment.
- Calcul du temps écoulé entre chaque point pour déterminer la vitesse et l'allure.

Sortie : `DataFrame` contenant les informations suivantes pour chaque segment : latitude, longitude, élévation, distance, différence d'altitude, pente, temps écoulé, allure.

5.2 Étape 2 : Application du Modèle de Minetti

Entrée : Pente de chaque segment.

Traitement :

- Application de la fonction de coût énergétique de Minetti pour chaque pente.
- Calcul du coût énergétique relatif par rapport à une surface plane.
- Ajustement de l'allure en fonction du coût énergétique.

Sortie : Allure ajustée pour chaque segment, reflétant l'effort requis en fonction de la pente.

5.3 Étape 3 : Intégration des Facteurs Environnementaux via Apprentissage Automatique

Entrée : Coordonnées moyennes du parcours pour déterminer la localisation.

Traitement :

- Requête à l'API OpenWeatherMap pour obtenir les données météorologiques (température, humidité).
- Application de pénalités sur l'allure ajustée en fonction de la température ($>20^{\circ}\text{C}$) et de l'humidité ($>60\%$).

Sortie : Allure ajustée en fonction des conditions environnementales pour chaque segment.

5.4 Étape 4 : Modélisation par Apprentissage Automatique

Entrée : Données du parcours (pente, allure, etc.) et données environnementales (température, humidité, etc.).

Traitement :

- Préparation des données pour l'entraînement du modèle d'apprentissage automatique.
- Entraînement d'un modèle (par exemple, XGBoost) pour prédire l'allure en fonction des facteurs environnementaux.
- Validation du modèle sur un ensemble de données de test.

Sortie : Allure prédite pour chaque segment en fonction des conditions environnementales.

5.5 Étape 5 : Calcul des Temps Totaux

Entrée : Allures (originale, ajustée pour le terrain, ajustée pour l'environnement) pour chaque segment.

Traitement :

- Somme des allures pour chaque type d'ajustement.
- Conversion en temps total de course.

Sortie :

- Temps de course original (sans ajustement).
- Temps de course ajusté pour le terrain.
- Temps de course ajusté pour l'environnement.

5.6 Étape 6 : Comparaison des Deux Approches

Critères de comparaison :

- **Précision des prédictions :** Comparaison des temps de course ajustés obtenus par chaque méthode avec les performances réelles des coureurs.
- **Complexité de mise en œuvre :** Évaluation de la facilité d'implémentation et des ressources nécessaires pour chaque approche.
- **Adaptabilité :** Capacité de chaque méthode à s'adapter à différents types de parcours et conditions environnementales.
- **Transparence et interprétabilité :** Degré de compréhension des mécanismes internes de chaque modèle.

6 Analyse et Comparaison des Résultats

Les deux approches implémentées — le modèle de Minetti basé sur des formules physiologiques, et le modèle d'apprentissage automatique XGBoost — ont permis d'obtenir des estimations de l'allure corrigée en fonction du dénivelé et des conditions environnementales.

6.1 Résumé des Temps Obtenus

- **Temps original (plat) :** 1143.58 minutes
- **Temps ajusté par le modèle de Minetti (dénivelé uniquement) :** 1235.08 minutes
- **Temps ajusté par Minetti + météo :** 1261.02 minutes
- **Temps prédit par le modèle ML (XGBoost) :** 1263.57 minutes

6.2 Écart entre les approches

- L'ajout du facteur météorologique augmente logiquement le temps de course corrigé d'environ **+2%** par rapport à la version dénivelée seule.
- Le modèle XGBoost prédit un temps **légèrement supérieur** à celui issu du modèle Minetti enrichi (1263.57 min vs 1261.02 min), ce qui suggère que le modèle a bien capté l'impact combiné des variables environnementales.
- L'erreur quadratique moyenne (RMSE) du modèle XGBoost est de **40.45**, ce qui montre une précision acceptable compte tenu de la simplicité des données simulées.

6.3 Interprétation et Choix Méthodologique

- Le **modèle de Minetti** est simple, transparent et efficace pour corriger l'allure en fonction de la pente. Il constitue une excellente base pour une première estimation robuste.
- Le **modèle d'apprentissage automatique** offre une personnalisation et une extensibilité accrues, surtout lorsqu'on dispose d'un volume plus important de données réelles.
- Dans un contexte de production, la **combinaison des deux approches** est recommandée : Minetti pour l'interprétabilité immédiate, et ML pour la personnalisation à grande échelle.

Conclusion : Cette analyse comparative permettra de déterminer l'approche la plus appropriée pour une utilisation en production, en tenant compte des avantages et des inconvénients de chacune.

Références Bibliographiques

1. Minetti, A. E., Moia, C., Roi, G. S., Susta, D., & Ferretti, G. (2002). Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of Applied Physiology*, 93(3), 1039–1046. <https://doi.org/10.1152/jappphysiol.01177.2001>
2. Weiss, K., Valero, D., Villiger, E., Scheer, V., Thuany, M., Aidar, F. J., ... & Knechtle, B. (2024). Associations between environmental factors and running performance : An observational study of the Berlin Marathon. *PLOS ONE*, 19(10), e0312097. <https://doi.org/10.1371/journal.pone.0312097>
3. Lemire, M., Falbriard, M., Aminian, K., Millet, G. P., & Meyer, F. (2021). Level, Uphill, and Downhill Running Economy Values Are Correlated Except on Steep Slopes. *Frontiers in Physiology*, 12, 697315. <https://doi.org/10.3389/fphys.2021.697315>
4. Knechtle, B., Nikolaidis, P. T., & Rosemann, T. (2021). The Influence of Environmental Conditions on Pacing in Age Group Marathoners Competing in the “New York City Marathon”. *Frontiers in Physiology*, 13, 842935. <https://doi.org/10.3389/fphys.2022.842935>