

Projet Hackathon Vélib's: recherche d'une répartition matinale optimale des vélos

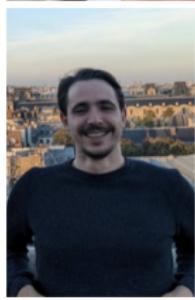
Mehdi Boussâa, Diala Hawat,

Alexander Reisach, Mariem Abaach, Raphaël Lachièze-Rey,

22.12.2023

Laboratoire MAP5





Problème: la régulation des stations

Régulation des stations

- Il s'agit de corriger les déséquilibres dans la répartition des vélos sur le réseau générés par les usagers.

Régulation des stations

- Il s'agit de corriger les déséquilibres dans la répartition des vélos sur le réseau générés par les usagers.



(a) Station auto-équilibrée



(b) Station auto-équilibrée

Régulation des stations

- Il s'agit de corriger les déséquilibres dans la répartition des vélos sur le réseau générés par les usagers.



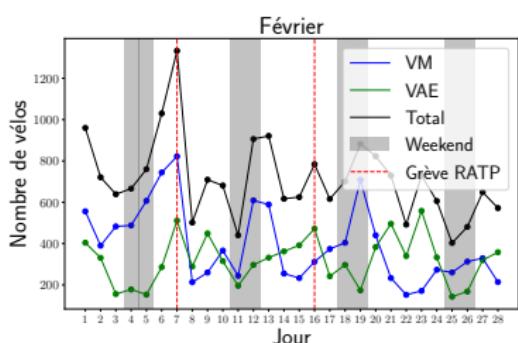
(a) Station remplie le soir



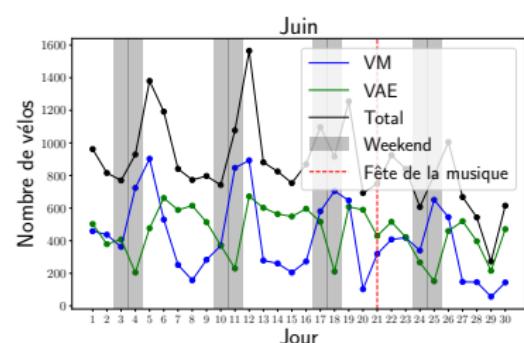
(b) Station vide le soir

Régulation des stations

- Il s'agit de corriger les déséquilibres dans la répartition des vélos sur le réseau générés par les usagers.
- Nombre moyen de vélos utilisés dans la régulation : 714 en février et 873 en juin.



(a) Régulation faite en février



(b) Régulation faite en juin

Régulation des stations

- Il s'agit de corriger les déséquilibres dans la répartition des vélos sur le réseau générés par les usagers.
- Nombre moyen de vélos utilisés dans la régulation : 714 en février et 873 en juin.

Comment peut-on rendre la régulation la plus efficace possible?

Données fournies par Smovengo

Contexte

- 1471 stations sur un territoire de 450 km^2 .
 - ≥ 15000 vélos (mécaniques et électriques).
 - Chaque seconde 5 vélos pris en station en pic d'activité.
- 5 jeux de données fournies sur 2 mois : février et juin.
 - Variabilité inter mensuelle dûe à la météo.
 - Influence d'évenement rares (grèves, jour fériés).

Données

1. Courses usagers.

| Numéro de Vélo | Assistance électrique | ID utilisateur | Date départ | Date arrivée | Distance parcourue en mètres | Durée en secondes | Vitesse maximum | Nom station départ | Latitude station départ | Longitude station départ | Statut station départ | Nom station arrivée |
|----------------|-----------------------|----------------|---------------------|---------------------|------------------------------|-------------------|-----------------|--------------------------|-------------------------|--------------------------|-----------------------|---|
| 2041 | True | 71709 | 2023-02-02 23:52:49 | 2023-02-03 00:06:26 | 4109 | 812 | 31 | Damrémont - Ordener | 48.893400 | 2.336267 | NORMALE | Laumi  re - Petit |
| 2041 | True | 13661 | 2023-02-03 08:17:17 | 2023-02-03 08:45:02 | 7591 | 1659 | 26 | Laumi  re - Petit | 48.884674 | 2.380039 | NORMALE | Gabriel Lam   |
| 2041 | True | 44481 | 2023-02-03 08:51:06 | 2023-02-03 09:11:07 | 3524 | 1196 | 26 | Gabriel Lam   | 48.835121 | 2.385535 | NORMALE | Gabriel Lam   |
| 2041 | True | 93548 | 2023-02-03 12:10:23 | 2023-02-03 12:19:19 | 1735 | 531 | 28 | Gabriel Lam   | 48.835121 | 2.385535 | NORMALE | Quai Panhard et Levassor |
| 2041 | True | 1601 | 2023-02-03 17:36:32 | 2023-02-03 18:15:03 | 10980 | 2305 | 29 | Quai Panhard et Levassor | 48.828375 | 2.384274 | NORMALE | Andr   Mauvois - Joseph et Marie Hackin |

Données

2. Historique remplissage des stations.

| Code station | Nom station | longitude | latitude | Date mise à jour | VM disponibles | VAE disponibles | VM indisponibles | VAE indisponibles | Nombre de diapasons disponibles |
|--------------|-------------------------------|-----------|-----------|---------------------|----------------|-----------------|------------------|-------------------|---------------------------------|
| 1001 | Quai de l'Horloge - Pont Neuf | 2.341798 | 48.857059 | 2023-02-01 00:00:00 | 1 | 0 | 0 | 0 | 16 |
| 1001 | Quai de l'Horloge - Pont Neuf | 2.341798 | 48.857059 | 2023-02-01 00:02:00 | 1 | 0 | 1 | 0 | 15 |
| 1001 | Quai de l'Horloge - Pont Neuf | 2.341798 | 48.857059 | 2023-02-01 03:25:00 | 1 | 1 | 1 | 0 | 14 |
| 1001 | Quai de l'Horloge - Pont Neuf | 2.341798 | 48.857059 | 2023-02-01 06:15:00 | 1 | 0 | 1 | 0 | 15 |
| 1001 | Quai de l'Horloge - Pont Neuf | 2.341798 | 48.857059 | 2023-02-01 06:29:00 | 1 | 1 | 1 | 0 | 14 |

Données

3. Régulation mouvements entre stations.

| Nom station prise | Latitude station prise | Longitude station prise | Nom station dépose | Latitude station dépose | Longitude station dépose | Total | VAE | VM |
|--|------------------------|-------------------------|--|-------------------------|--------------------------|-------|-----|----|
| Javel - André Citroën | 48.846276 | 2.278637 | Edgar Quinet - Gaité | 48.841161 | 2.324595 | 13 | 2 | 11 |
| Place de la Montagne du Goulet | 48.844341 | 2.277384 | Charles de Gaulle - Plateau de Vanves | 48.822331 | 2.298572 | 12 | 2 | 10 |
| Eugène Oudiné - Dessous Des Berges | 48.825642 | 2.375023 | Bobillot - Paulin Mery | 48.829900 | 2.354336 | 9 | 1 | 8 |
| La Jarry - Place Diderot | 48.849727 | 2.452014 | Fabre d'Eglantine - Place de la Nation | 48.847112 | 2.395620 | 11 | 4 | 7 |
| Gabriel Péri - Danielle Casanova | 48.790114 | 2.402578 | Cours de Vincennes - Soult | 48.846919 | 2.410522 | 13 | 7 | 6 |
| Place Marcel Cachin | 48.814114 | 2.390908 | François Debergue - Croix de Chavaux | 48.857422 | 2.437578 | 3 | 0 | 3 |
| Place Marcel Cachin | 48.814114 | 2.390908 | Louis Lumière. | 48.862162 | 2.411126 | 6 | 0 | 6 |
| Gouthière - Poterne des Peupliers | 48.820507 | 2.351342 | Château - Place de Catalogne | 48.836806 | 2.318529 | 9 | 0 | 9 |
| Aristide Briand - Place de la Résistance | 48.821242 | 2.251100 | Terroirs de France - Lheureux | 48.832120 | 2.388618 | 6 | 0 | 6 |
| Censier - Jardin des Plantes | 48.841072 | 2.355475 | Place de la Nation - Picpus | 48.846822 | 2.400113 | 13 | 1 | 12 |
| David Weill - Parc Montsouris | 48.821092 | 2.333726 | Boulard - Daguerre | 48.834465 | 2.329339 | 9 | 4 | 5 |
| Javel - André Citroën | 48.846276 | 2.278637 | Château - Place de Cataloane | 48.836806 | 2.318529 | 7 | 0 | 7 |

Données

4. Régulation mouvement ateliers.

- Mouvement atelier-vers-station.

| Nom station dépose | Latitude station dépose | Longitude station dépose | Total | VAE | VM |
|---|-------------------------|--------------------------|-------|-----|----|
| Gare du Nord - Hôpital Lariboisière | 48.881949 | 2.352339 | 4 | 0 | 4 |
| Place d'Italie - Vincent Auriol | 48.831482 | 2.356755 | 12 | 10 | 2 |
| Station Paris Basketball - Stade Georges Carpe... | 48.820482 | 2.366832 | 9 | 6 | 3 |
| Raymond Poincaré - Place Victor Hugo | 48.870306 | 2.285075 | 9 | 9 | 0 |
| Porte Molitor | 48.845248 | 2.256902 | 9 | 0 | 9 |
| Grande Armée - Brunel | 48.876116 | 2.288124 | 14 | 10 | 4 |
| Diane Arbus | 48.898801 | 2.352402 | 13 | 13 | 0 |
| Frères Flavien - Porte des Lilas | 48.878099 | 2.411046 | 14 | 10 | 4 |
| Paul Signac - Aristide Briand | 48.869503 | 2.450226 | 12 | 7 | 5 |
| Porte des Lilas - Faidherbe | 48.878619 | 2.412033 | 30 | 15 | 15 |
| Le Goff - Gay-Lussac | 48.846102 | 2.341221 | 14 | 9 | 5 |
| Hoche - Tilsitt | 48.874836 | 2.297132 | 9 | 3 | 6 |

Données

4. Régulation mouvement ateliers.

- Mouvement atelier-vers-station.
- Mouvement station-vers-atelier.

| Nom station prise | Latitude station prise | Longitude station prise | Total | VAE | VM |
|---------------------------------------|------------------------|-------------------------|-------|-----|-----|
| Alexandre Dumas - Voltaire | 48.852889 | 2.389148 | 3 | 3 | 0 |
| Amelot - Saint-Sébastien | 48.861142 | 2.367800 | 2 | 2 | 0 |
| André Theuriet - Blagis | 48.779773 | 2.313053 | 1 | 1 | 0 |
| Archives - Rivoli | 48.857347 | 2.353977 | 1 | 0 | 1 |
| Arcole - Notre-Dame | 48.853884 | 2.349380 | 2 | 0 | 2 |
| ... | ... | ... | ... | ... | ... |
| Vincennes - Douy-Delcupe | 48.854937 | 2.436049 | 1 | 1 | 0 |
| Vincent Auriol - Louise Weiss | 48.834002 | 2.370389 | 1 | 1 | 0 |
| Vistule - Choisy | 48.823655 | 2.361572 | 1 | 1 | 0 |
| Vivaldi - Antoine-Julien Hénard | 48.841924 | 2.389729 | 2 | 1 | 1 |
| Westermeyer - Paul Vaillant-Couturier | 48.819116 | 2.396664 | 1 | 1 | 0 |

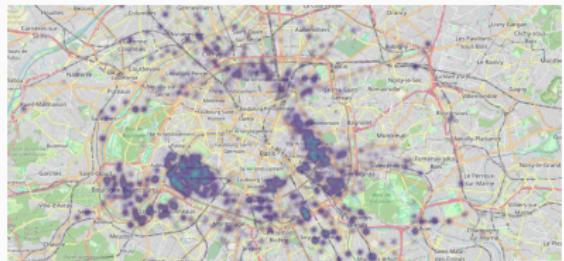
Données

5. Signalement vélo abandonnés.

| Date de signalement | Numéro de Vélo | Localisation |
|-------------------------|----------------|---|
| 2023-02-01 08:37:43.285 | 57050 | 23 Voie Br/15, 75015 Paris, France |
| 2023-02-01 09:01:40.695 | 5803 | 14 Rue De Dunkerque, 75010 Paris, France |
| 2023-02-01 11:08:17.641 | 62660 | 37 Rue Clavel 75019 Paris |
| 2023-02-01 12:34:14.555 | 31700 | 4 Rue Francis Jammes, 75010 Paris, France |
| 2023-02-01 12:40:07.261 | 34111 | 128 Avenue De La République, 93300 Aubervillie... |
| ... | ... | ... |
| 2023-02-28 18:32:51.593 | 58407 | 20 Rue Bénard, 75014 Paris, France |
| 2023-02-28 21:00:47.310 | 52291 | 3 Avenue Saint-Joseph 92600 Asnières-sur-Seine |
| 2023-02-28 23:08:44.199 | 52865 | 40 Rue De La Fédération, 93100 Montreuil, France |
| 2023-02-28 23:10:41.466 | 60010 | Galerie De La Villette, 75019 Paris, France |
| 2023-02-28 23:51:49.985 | 31230 | 84 Rue Réaumur, 75002 Paris, France |

Analyse de données

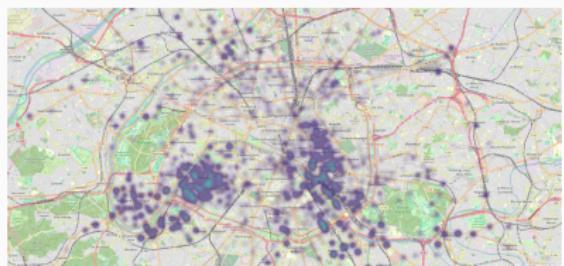
Analyse de données



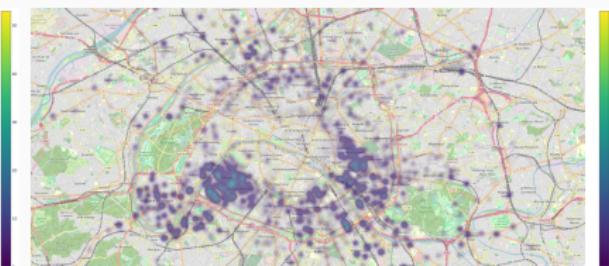
(a) Répartition des vélos à 6h le 16 juin



(b) Répartition des vélos à 20h le 16 fév

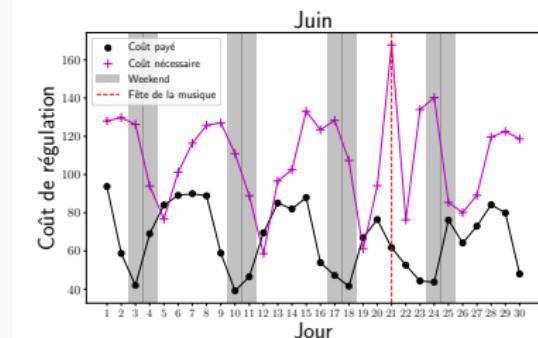


(c) Répartition des vélos à 1h le 17 juin

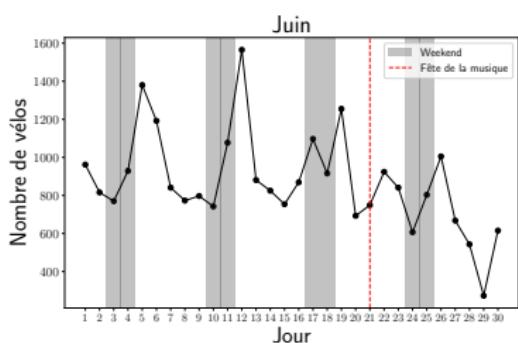


(d) Répartition des vélos à 6h le 17 juin

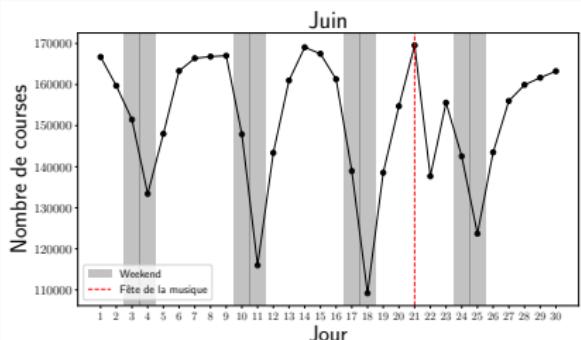
Analyse de données



(a) Coût de regulation

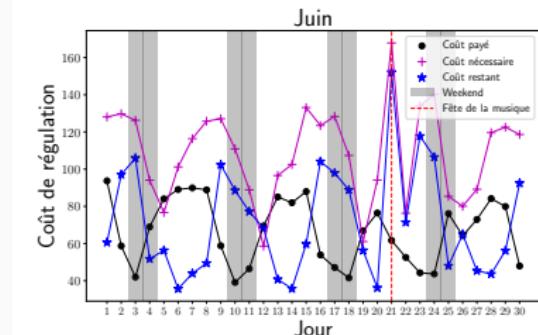


(b) Vélos utilisé dans la régulation

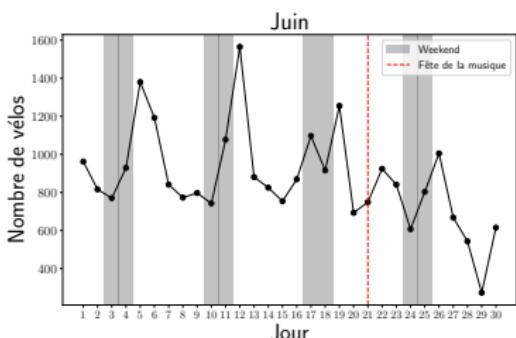


(c) Nombre de courses

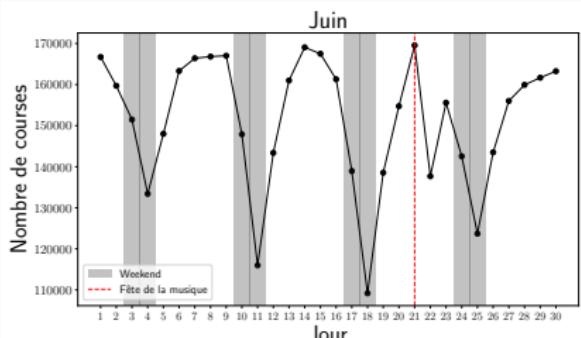
Analyse de données



(a) Coût de regulation

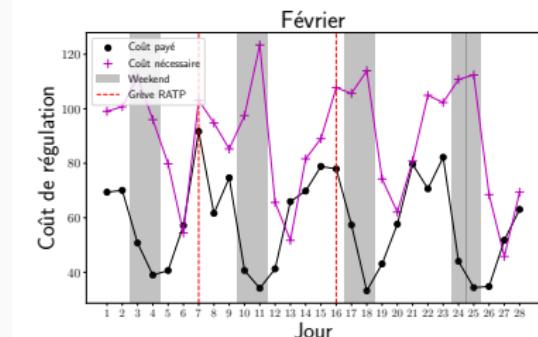


(b) Vélos utilisé dans la régulation

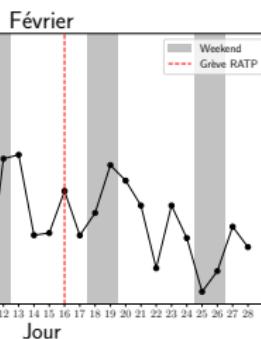


(c) Nombre de courses

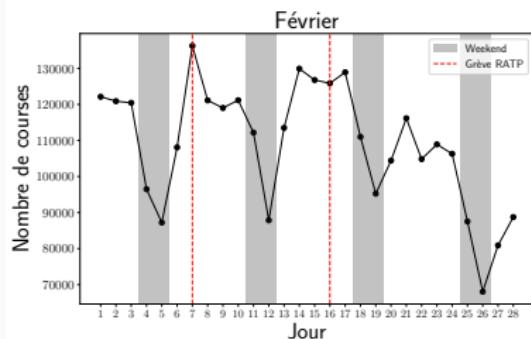
Analyse de données



(a) Coût de regulation

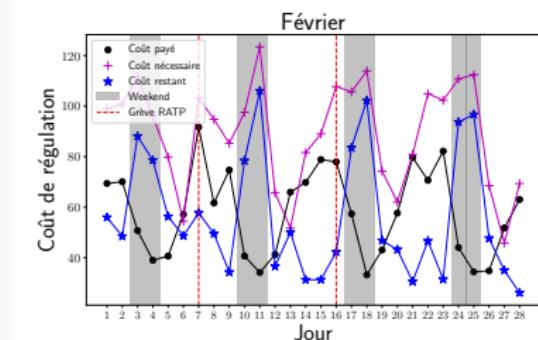


(b) Vélos utilisé dans la régulation

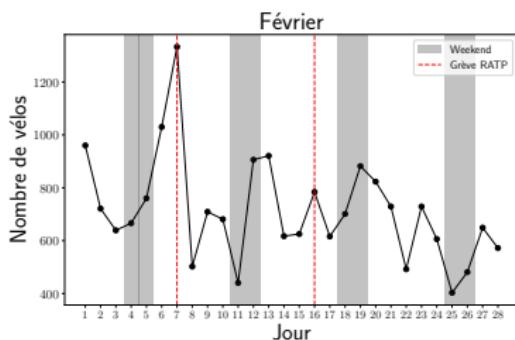


(c) Nombre de courses

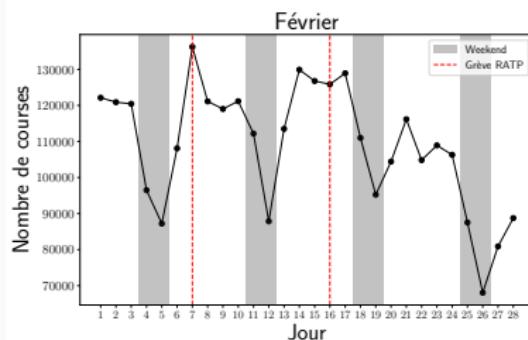
Analyse de données



(a) Coût de regulation



(b) Vélos utilisé dans la régulation



(c) Nombre de courses

Projet proposé

Objectifs

Proposer un outil mathématique améliorant la régulation en respectant ces critères :

- Réduire les coûts (carbone et frais).
- Maximiser la satisfaction des utilisateurs.
- Être réalisable et implémentable.

Idée

- Attribuer à une répartition du matin des vélos μ^m un score.
- Trouver une répartition matinale optimale μ_{opt}^m .

Outil algorithmique utilisé par l'équipe de régulation chaque soir pour prédire à partir de la répartition du soir μ^s la répartition matinale μ_{opt}^m qui maximise le score.

Optimiser la régulation des vélos pour une efficacité maximale

On définit le score d'une répartition matinale de vélos μ^m en prenant en compte:

- Le nombre d'utilisateurs satisfaits n .
- Le coût W demandé pour parvenir à μ^m à partir de μ^s .

Exemples:

$$s(\mu^m) = \frac{n(\mu^m)}{W_\lambda(\mu^s, \mu^m)} \quad \text{ou} \quad s(\mu^m) = n(\mu^m) - \lambda W(\mu^s, \mu^m).$$

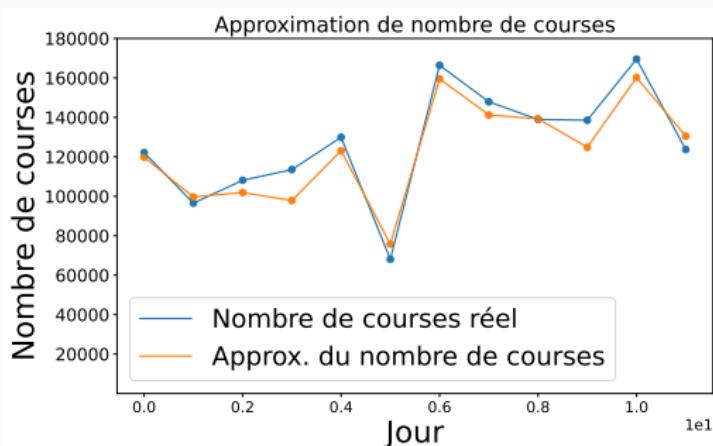
- $n(\mu^m)$: nombre de courses pour une répartition matinale μ^m .
- $W(\mu^s, \mu^m)$: coût minimal pour transporter μ^s vers μ^m .
- λ : Paramètre d'équilibre positif.

Exemple jouet

Maximisation du score

Exemple jouet:

- $n(\mu^m)$: un réseau de neurones à trois couches entièrement connecté, avec une fonction d'activation *Relu*.



Maximisation du score

Exemple jouet:

- $n(\mu^m)$: un réseau de neurones à trois couches entièrement connecté, avec une fonction d'activation *Relu*.
- $W(\mu^s, \mu^m)$: distance de Wasserstein d'ordre 2 regularisée

$$\min_{\pi \in \Pi(\mu^s, \mu^m)} \sum_{i,j} c_{ij} \pi_{ij} - \varepsilon H(\pi),$$

Où $H(\pi) = - \sum_{i,j} \pi_{ij} \log(\pi_{ij})$ et $\varepsilon > 0$.

Maximisation du score

Exemple jouet:

- $n(\mu^m)$: un réseau de neurones à trois couches entièrement connecté, avec une fonction d'activation *Relu*.
- $W(\mu^s, \mu^m)$: distance de Wasserstein d'ordre 2 regularisée

On va chercher μ_{opt}^m tel que, pour $\lambda > 0$

$$\mu_{\text{opt}}^m \in \operatorname{argmax} s(\mu^m) = n(\mu^m) - \lambda W(\mu^s, \mu^m).$$

Optimisation via Adam :

$$\mu_i = \mu_i + f(\nabla s(\mu_i)).$$

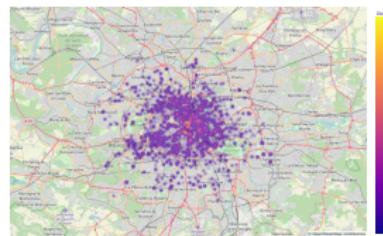
Quelques résultats



(a) Répartition du soir



(b) Répartition réelle du matin



(c) Répartition optimale du matin

Amélioration W : +6.5%, n : + 10%, 52% non déplacés (en W)

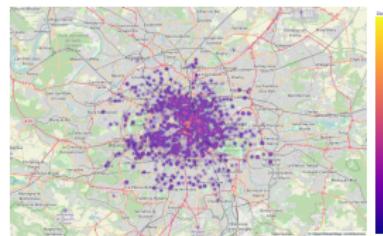
Quelques résultats



(a) Répartition du soir



(b) Répartition réelle du matin



(c) Répartition optimale du matin

Amélioration W : +6%, n : + 18%, 51% non déplacés (en W)

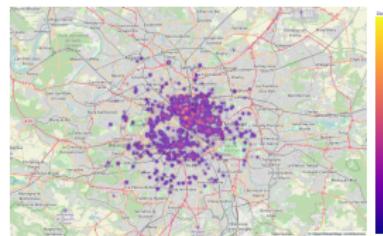
Quelques résultats



(a) Répartition du soir



(b) Répartition réelle du matin



(c) Répartition optimale du matin

Amélioration W : -100%, n : + 150%, 39% non déplacés (en W)

Résumé, avantages, perspectives

Résumé, avantages, perspectives

Résumé :

- Fonction score évaluant l'efficacité d'une répartition matinale.
- Outil algorithmique visant à identifier la répartition matinale optimale maximisant ce score.

Résumé, avantages, perspectives

Résumé :

- Fonction score évaluant l'efficacité d'une répartition matinale.
- Outil algorithmique visant à identifier la répartition matinale optimale maximisant ce score.

Avantages :

- Réduire les coûts de régulation.
- Réduire l'empreinte carbone de la régulation.
- Augmentation de la satisfaction des utilisateurs.

Résumé, avantages, perspectives

Résumé :

- Fonction score évaluant l'efficacité d'une répartition matinale.
- Outil algorithmique visant à identifier la répartition matinale optimale maximisant ce score.

Avantages :

- Réduire les coûts de régulation.
- Réduire l'empreinte carbone de la régulation.
- Augmentation de la satisfaction des utilisateurs.

Perspectives :

- Intégrer les contraintes de l'équipe Vélib's dans le modèle.
- Étude plus approfondie avec plus de données.

Merci pour votre attention !
