

# Valeurs de l'information et performances algorithmiques pour des problèmes multi-périodes et stochastiques

Th. Pironet

HEC-Ecole de gestion de l'Université de Liège

## Table des matières :

- Message principal
- Transport : problèmes classiques
- Multi-périodes : informations, décisions et stochasticité
- Optimisation : scénario, simulation et bornes
- Valeurs de l'information : définitions et diagrammes
- Performances algorithmiques : 6 + 1 méthodologies
- Résumé
- 2 applications
- Conclusions

# Message principal

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Résumé :

**Comment mesurer** ce que nous perdons lorsque nous décidons

- ① sans connaître tout jusqu'à la **fin des temps**,
- ② sans connaître tout avec **certitude**,
- ③ et par conséquent de **manière répétitive**,

alors que l'on pourrait décider **une fois** avec **omniscience** ?

## Informations et décisions :

**Manques d'infos quantitativement et qualitativement**

# Problèmes classiques de transport

## Informations

- Réseau (liens, temps/coûts de transport)
- Flotte de véhicules (taille, capacité, localisations)
- Demande (quantités, localisations, fenêtres de temps)

## Décisions

Quels véhicules pour quels clients ? Quand ? Quels parcours ?  
=> Sélection, routage, chargement, ordonnancement (Entiers)



GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

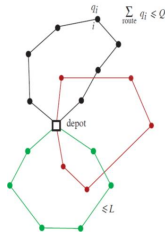
# Problèmes classiques de transport

## Formulation

**Objectif** Minimiser les coûts ou maximiser les profits

### Contraintes

- Satisfaire la demande
- Tournées avec départs et arrivées au(x) dépôt(s),
- Respect des capacités
- Respect des fenêtres de temps
- etc...



GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

# Informations Multi-périodes

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Modèles dans la littérature : Distribution journalière

Problèmes étudiés : solution **unique** et données **déterministes**

Méthodes exactes ou heuristiques

Modifications : **Informations** et **Décisions** multi-périodes

## Informations multi-périodes

### Horaires de bus et train :

Demande annuelle par heure/jour/semaine (congés scolaires)

=> une solution globale qui couvre chaque période

### Horaires périodiques de livraison/collecte :

Demande par semaine stable à répartir

=> une solution globale qui couvre 5-7 périodes et à répéter

Stabilités : demande des clients et des localisations de la flotte

Informations multi-périodes **///#//** décisions répétitives, **mais**

# Décisions multi-périodes

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Décisions multi-périodes

### Décisions consécutives interreliées

- demande non connue (localisations, quantités et moments)
- variations des localisations initiales des véhicules

Traditionnellement, transport longues distances

Tendance : infos en ligne (Mob., GPS, e-commerce, trafic)

=> Décisions successives lors de la période en cours  $t$

Fréquences : jour après jour, heure en heure, en continu...

## Décisions successives

$T$  Solutions pour  $t = 1, 2, \dots, T$  => **Une Politique**  $\pi$

Objectif : économique, optimiser l'espérance  $\mu_\pi$

Problème de **faisabilité**, i.e., pas de pénalités infinies

# Informations multi-périodes et stochastiques

Données : **D**écision, **D**éterministe, **S**tochastique



Dynamique du système de décision sur les horizons :

- 1 Actions en  $t$  et ajout de leurs valeurs à  $\mu_\pi$
- 2 Sorties d'informations
- 3 On déroule 1 période, mise à jour  $t \rightarrow t + 1$
- 4 Entrées d'informations
  - 1 Du **Stochastique** devient **Déterministe** :  
 $t + RH + 1 \rightarrow t + RH$
  - 2 Une période de la fin d'horizon devient **Stochastique** :  
 $t + H + 1 \rightarrow t + H$
- 5 Retour à l'étape 1 avec  $t \rightarrow t + 1$



# Optimisation multi-périodes et stochastiques

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

**Optimisation**

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## 2 causes de sous-optimalité

La succession des décisions optimales (une politique)  
 $\leq$   
la séquence optimale de décision (la solution)

① la politique basée sur **l'horizon fini**

$\leq$   
la solution basée **l'horizon infini**

② la politique  $\mu_\pi$  basée sur **l'info stochastique**

$\leq^*$   
la politique basée sur **l'info déterministe**

## 2 tentatives de réponse

① Comment mesurer les valeurs de l'information ?

② Comment mesurer les performances algorithmiques ?

# Optimisation : Scénario

## Scénario :

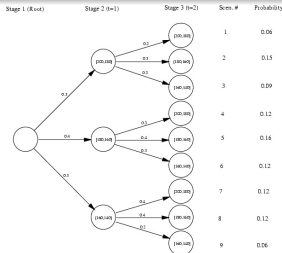
Une réalisation **déterministe** des paramètres **stochastiques**  
e.g., des distributions discrètes  $\Rightarrow$  **un nombre fini** de scénarios

$t$  |  $t+1, \dots, t+RH$  |  $t+RH+1, \dots, t+H$

## Déterministe

## Modèle déterministe équivalent : Arbre de scénarios

**Tous les** scénarios **liés** par des contraintes de non-anticipativité  
Égalité des décisions en  $t$  pour tous les scénarios



# Optimisation : Arbre de scénarios

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu  
Message  
Transport  
Multi-périodes

Optimisation

Scénario  
Bornes

Valeurs infos

Diagrammes

Performances

Mesures  
Méthodologie

Résumé

Applications

Conclusions

## En théorie :

La solution du modèle "arbre de scénarios" est optimale

## En pratique : Modèles de grandes tailles

Exemple : un chargement est disponible ou non, très simple  
N chargements :  $2^N$  scénarios  $\implies 2^{20} \cong 10^6$ ,  $2^{30} \cong 10^9$

$\implies$  **"Insoluble"** : temps, tailles (mémoire et calculs),  
formulation plus complexe que pour un seul scénario

$\implies$  Approximation  $\mu_{\pi^*}$

$\implies$  Simulation sur  $N$  scénarios pour valider  
Valeur espérée :  $E...$

# Simulation : Bornes

GT2L Lille  
30/3

Th. Pironet



## Bornes : Scénarios avec informations déterministes

- 1 **Politique** Myope (utilisable) pour  $t : O_1^*$
- 2 **Politique** Info partiellement révélée pour  $RH : O_{RH}^*$
- 3 **Politique** Oracle fini (a-posteriori) pour  $H : O_H^*$
- 4 **Solution** Oracle "infini" (a-posteriori) pour  $O_T^*$  ( $T \rightarrow \infty$ )

## Scénario espéré : équivalent déterministe "représentatif"

Réalisation de l'espérance de la distribution :  $EVS$

## Bornes sur la politique, on "Espère" que

Si Maximisation :  $O_T^* \geq O_H^* \geq \mu_{\pi^*} \geq EVS \geq O_{RH}^* \geq O_1^*$

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

# Valeurs des informations (Max)

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

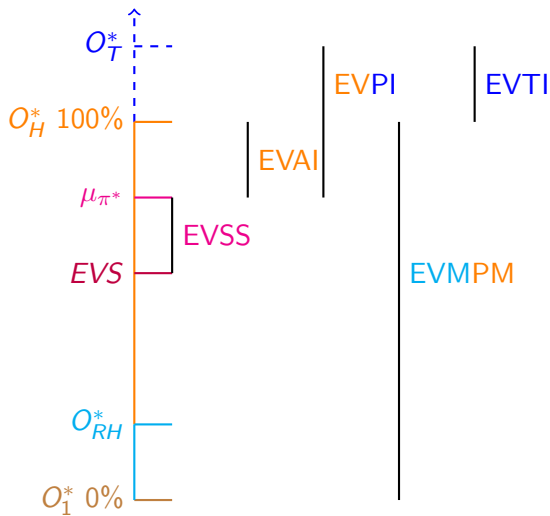
Mesures

Méthodologie

Résumé

Applications

Conclusions



# Valeurs des informations : exemples

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

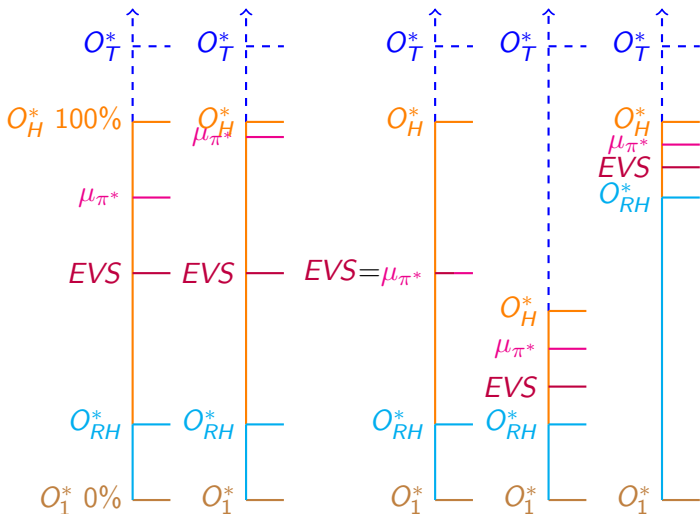
Mesures

Méthodologie

Résumé

Applications

Conclusions



# Valeurs d'information : Pour quoi ?

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Valeurs d'information "Grandes" :

Améliorations potentielles : Coûts / Bénéfices (%)

- 1 Changer  $RH$  i.e., confirmation de commandes ;
- 2 Changer  $H$  i.e., méthodes de prévisions ;
- 3 Changer  $\mu_{\pi^*}$  i.e., méthodes d'optimisation

## Valeurs d'information "Petites" :

Utiliser un modèle plus simple :  $O_T^* \cong EVS \cong O_{RH}^* \cong O_1^*$

## Valeurs d'information "Petites" et "Grandes" ?

Comment évaluer l'écart entre deux valeurs ?

Valeurs proches...Pourcentages...

=> **Mesure des performances algorithmiques**

## Liste non-exhaustive

- 1 Valeur de l'information "sous heuristique" (Mitrovic Minic et al. 2004),
- 2 Moyennes et ratios (Regan et al. 1998),
- 3 Analyse comparative et rapport d'approximation (Angelelli et al. 2007),
- 4 Pourcentage à la solution a-posteriori ou bornes (Spivey et Powell 2004),
- 5 Compromis entre moyenne et variance (List et al. 2003),
- 6 Test statistique de comparaison sur un échantillon apparié (Tjokroamidjojo et al. 2006)



# Performances algorithmiques 1

## Valeur de l'information "sous heuristique"

Pour un problème multi-périodes déterministe

**Valeur de l'information :**

$$V = \frac{O_T^* - O_{H=RH}^*}{O_{H=RH}^*}$$

Objectif : mesurer ce qui est perdu à la fin de l'horizon  $\cong$  *EVTI*  
**Option "sous heuristique"**

**Approximations  $\mathcal{H}$  des valeurs :**

$$V = \frac{O_T^{\mathcal{H}} - O_{H=RH}^{\mathcal{H}}}{O_{H=RH}^{\mathcal{H}}}$$

- Modèle multi-périodes approché
- Méthodes algorithmiques exactes ou approchées
- Si stochastique, nombre de scénarios de validation

*EVTI* et Approximations des valeurs

## Moyennes et ratios

Affectation de chargements à une flotte

- Règle : le camion le plus proche
- Heuristiques avec un horizon multi-périodes
- En-ligne

### Résultats :

Moyennes et écarts-type :

- du nombre de trajets à vide et chargés
- des distances parcourues à vide et chargés
- du temps moyen d'attente pour effectuer le service

Ratio du temps chargé, profit opérationnel espéré, ratio des enlèvements en retard

**Moyennes et surtout écarts-type, ratios... ? Unités !**

# Performances algorithmiques 3

## Analyse comparative et rapport d'approximation

Rapport d'approximation de  $\mathcal{H}$  sur une instance  $I$   
Valeur  $v(\mathcal{H}, I)$  de l'heuristique et valeur optimale  $v^*(I)$   
Pour une minimisation<sup>(1)</sup> ou maximisation<sup>(2)</sup>

$$r_{\mathcal{H}} = \max_I \left( \frac{v(\mathcal{H}, I)}{v^*(I)} \text{ (1)}, \frac{v^*(I)}{v(\mathcal{H}, I)} \text{ (2)} \right) \geq 1$$

Pour un problème multi-périodes déterministe  
Heuristique : information révélée sur  $H$ , "en-ligne"  $\mathcal{H}_H$   
Optimal : information révélée sur  $T$ , "hors-ligne"  $O_T^*$

$$r_{\mathcal{H}_H} = \max_I \left( \frac{v(\mathcal{H}_H, I)}{O_T^*} \text{ (1)}, \frac{O_T^*}{v(\mathcal{H}_H, I)} \text{ (2)} \right)$$

**Performance : pire théorique ou espérée en pratique**

# Performances algorithmiques 4

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Pourcentage à la solution a-posteriori ou bornes

e.g. : problème de maximisation :

- $O_T^* = 100\%$
- $O_1^* > 0\%$  (de 10 à 60%) **Autre pourcentage inférieur**
- $\mu_{\pi^*} = 50\%, 75\%$  ou  $91\%$ ? Est-ce bon ?

Exemples :

$$O_1^* = 10\%; \mu_{\pi^*} = 91\%$$

$$O_1^* = 90\%; \mu_{\pi^*} = 91\%$$

**Les pourcentages sont juste indicatifs**

**Importance des bornes**

**Performance avec l'information accessible ou parfaite**

**Normalisation de l'écart :  $O_1^* = 0\%$  et  $O_H^* = 100\%$**

**$EVMPM = 100\%$  à traduire en valeur économique**

## Compromis entre moyenne et variance

Performance en fonction de la robustesse

"Robustesse" : éviter les conséquences extrêmes

Parmi les politiques  $\pi_i$ , choisir  $\mu_{\pi^*}$  sauf si  $\sigma_{\pi^*} \geq \text{seuil}$

**Attention à l'écart-type, niveau de confiance**

## Test statistique de comparaison sur un échantillon apparié

Validation statistique de surclassement entre politiques  $\mu_{\pi(i)}$   
Avec une marge d'erreur de  $\alpha$ ,  
sur  $N$  scénarios de validation identiques,  
peut-on clamer **l'égalité** ou le **surclassement** :

$$H_0 : \mu_{\pi(1)} = \mu_{\pi(2)} \text{ vs } H_1 : \mu_{\pi(1)} > \mu_{\pi(2)} \text{ ou } H_1 : \mu_{\pi(1)} < \mu_{\pi(2)}$$

## Tests statistiques entre politiques

# Une proposition de méthodologie

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Validation statistique des performances :

- Calculer les bornes sur  $N \geq 30$  scénarios
- Vérifier la non "non-normalité" des résultats (Shapiro-Wilk)  
Echantillons "appariés" pour améliorer le taux de confiance
- Comparer :  $O_1^*$ ,  $O_{RH}^*$ ,  $O_H^*$  et  $O_T^*$  (Stop ou Encore)
- Calculer  $EVS$ , test S-W et comparer (Stop ou Encore)
- Calculer  $\pi_i$ , test S-W et comparer les  $\mu_{\pi_i}$  et les bornes
- Sélectionner  $\pi^*$

# Tests de non "non-normalité" et Z-test

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Test de Shapiro-Wilk : Web

<http://www2.cedarcrest.edu/academic/bio/hale/biostat/session25links/normality.html>

exemple : voir fichier excel

## Z-Test : Caractériser statistiquement la différence

- 1 Soit  $v(\mu_{\pi(1)}, i)$  et  $v(\mu_{\pi(2)}, i)$ , valeurs des politiques 1 et 2 sur le scénario  $i$  d'un échantillon apparié  $i = 1, \dots, N$
- 2 Evaluer les  $N$  différences  $d_i = v(\mu_{\pi(1)}, i) - v(\mu_{\pi(2)}, i)$
- 3 Evaluer la différence moyenne  $\bar{d} = \frac{1}{N} \sum_{i=1}^{i=N} d_i$
- 4 Evaluer la variance  $s_d^2 = \frac{1}{N} \sum_{i=1}^{i=N} (d_i - \bar{d})^2$
- 5 Evaluer la statistique pour le test  $z = \frac{\bar{d}}{s_d / \sqrt{(N-1)}}$
- 6 Pour une marge d'erreur  $\alpha$  (5% ou 1%), extraire la valeur critique de la table  $Z_{(1-\alpha)}$



## Z-Test : utilisation

- 1 Si  $|z| = \frac{|\bar{d}|}{s_d/\sqrt{(N-1)}} < Z_{(1-\alpha)}$ ,  
alors pour la marge d'erreur  $\alpha$ ,  
l'hypothèse  $H_0 : \mu_{\pi(1)} = \mu_{\pi(2)}$  ne peut pas être rejetée
- 2 Inversement, si  $|z| > Z_{(1-\alpha)}$   
alors pour la marge d'erreur  $\alpha$ 
  - l'hypothèse  $H_0$  peut être rejetée
  - et  $H_1 : \mu_{\pi(1)} > \mu_{\pi(2)}$  ou  $H_1 : \mu_{\pi(1)} < \mu_{\pi(2)}$  est acceptée
    - $\mu_{\pi(1)} > \mu_{\pi(2)}$  si  $z > 0$  càd  $\bar{d} > 0$
    - $\mu_{\pi(1)} < \mu_{\pi(2)}$  si  $z < 0$  càd  $\bar{d} < 0$

Exemple : voir fichier excel

# Tests complémentaires

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Analyse de sensibilité : $H$

Tester les performances en fonction de  $H$

## "Robustesse"

Tester les performances si mauvaise évaluation de la distribution

## "Biais" de démarrage et de fin d'horizon

Mise en route sur quelques périodes

Même période de mise en route

Valeur espérée moyenne sur  $T$  périodes  $\Rightarrow T$  long

## Méthodologie :

- Décrire les informations et horizons  $RH$ ,  $H$ ,  $T$  du problème multi-périodes
- Décrire le dynamisme des décisions  $t$
- Décrire la distribution si stochastique
- Générer un échantillon de scénarios ( $N \geq 30$ )
- Evaluer des bornes  $O_T^*$ ,  $O_H^*$ ,  $O_{RH}^*$ ,  $O_1^*$  et  $EVS$
- Evaluer les valeurs des informations et comparer statistiquement (Stop ou Encore)
- Générer des politiques, comparer et sélectionner  $\mu_{\pi^*}$
- Analyser la sensibilité à  $RH$  ou  $H$  et/ou la robustesse

# Applications en transport cas 1

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions



# Applications en transport cas 1

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Problème de chargement avec fenêtre de temps :

- Sous-arbre de scénarios (inutilisable : mémoire, temps)
- **EVS** très bon
- EVMPM significatif
- L'optimisation sur base du scénario "**Optimiste**", i.e. au plus tôt, surclasse les autres  $\pi_i$   
NB : pas toujours le meilleur, mais jamais surclassé!
- Rapide, modèle à un scénario
- Parfois statistiquement, "**Optimiste**" =  $O_H^*$
- **EVPI (EVTI 8.1% + EVAI 4.8%)**
- EVSS 6.1%, mais **EVS** surclassé

# Applications en transport cas 1

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

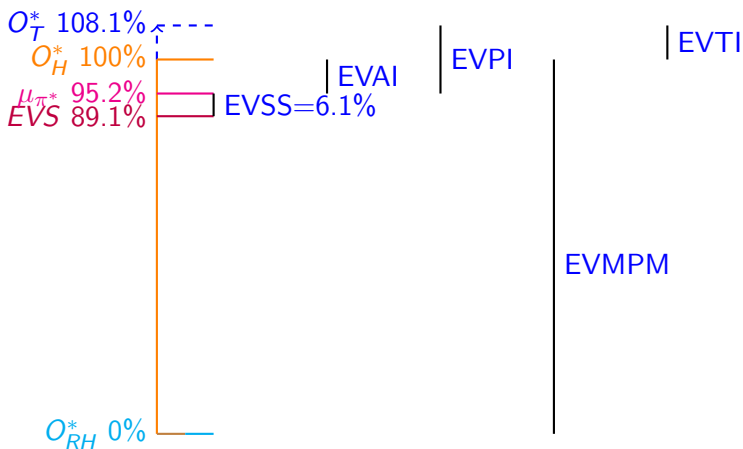
Mesures

Méthodologie

Résumé

Applications

Conclusions



# Applications en transport cas 2

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu  
Message  
Transport  
Multi-périodes

Optimisation

Scénario  
Bornes

Valeurs infos

Diagrammes

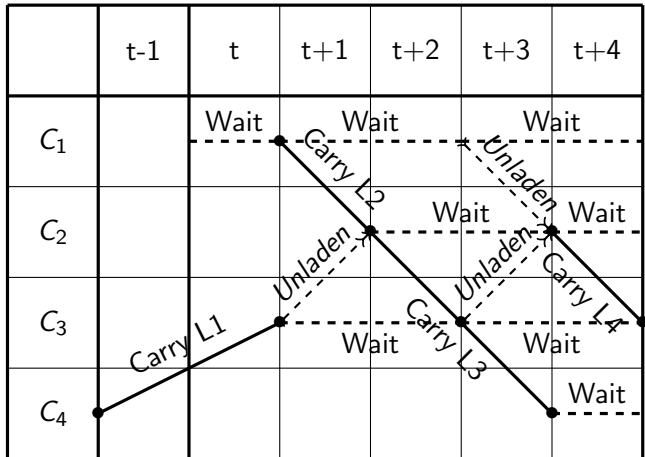
Performances

Mesures  
Méthodologie

Résumé

Applications

Conclusions



# Applications en transport cas 2

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Problème de routage de véhicules :

- *EVS* moyen 56.6%
- EVMPM significatif
- Sous-arbre de scénarios (10 à 50) surclasse les autres  $\pi_i$
- Les bornes sont rarement égalées statistiquement
- EVPI (EVTI 78.4%+ **EVAI 32.3%**)
- EVSS 11%, mais *EVS* surclassé



# Applications en transport cas 2

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

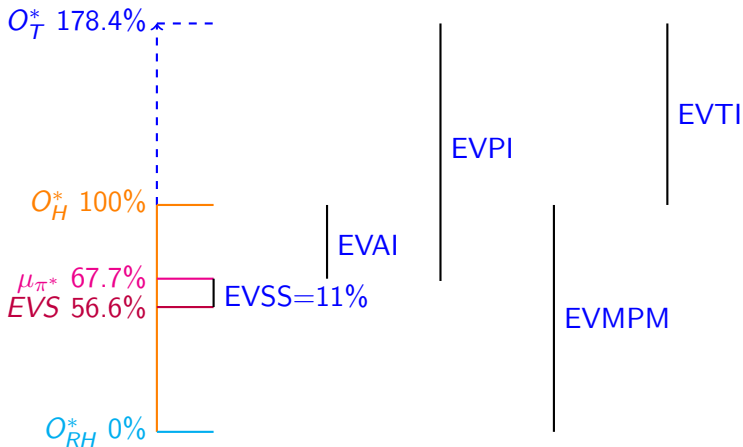
Mesures

Méthodologie

Résumé

Applications

Conclusions



# Conclusions

GT2L Lille  
30/3

Th. Pironet

Introduction

Contenu

Message

Transport

Multi-périodes

Optimisation

Scénario

Bornes

Valeurs infos

Diagrammes

Performances

Mesures

Méthodologie

Résumé

Applications

Conclusions

## Conclusions

- Intérêt pour les problèmes multi-périodes et stochastiques
- Description des informations disponibles et des horizons
- Description de la dynamique du processus de décision
- Bornes compréhensibles
- Définitions des valeurs de l'information
- Diagramme de visualisation et d'analyse des résultats
- Test statistique de mesure des performances
- Processus de décision pour développer des algorithmes

## Perspectives

**Algorithmes** pour problèmes multi-périodes stochastiques,  $\pi_i$  ?  
Prochain épisode...

Introduction

Contenu  
Message  
Transport  
Multi-périodes

Optimisation

Scénario  
Bornes

Valeurs infos

Diagrammes

Performances

Mesures  
Méthodologie

Résumé

Applications

Conclusions

Merci pour votre attention