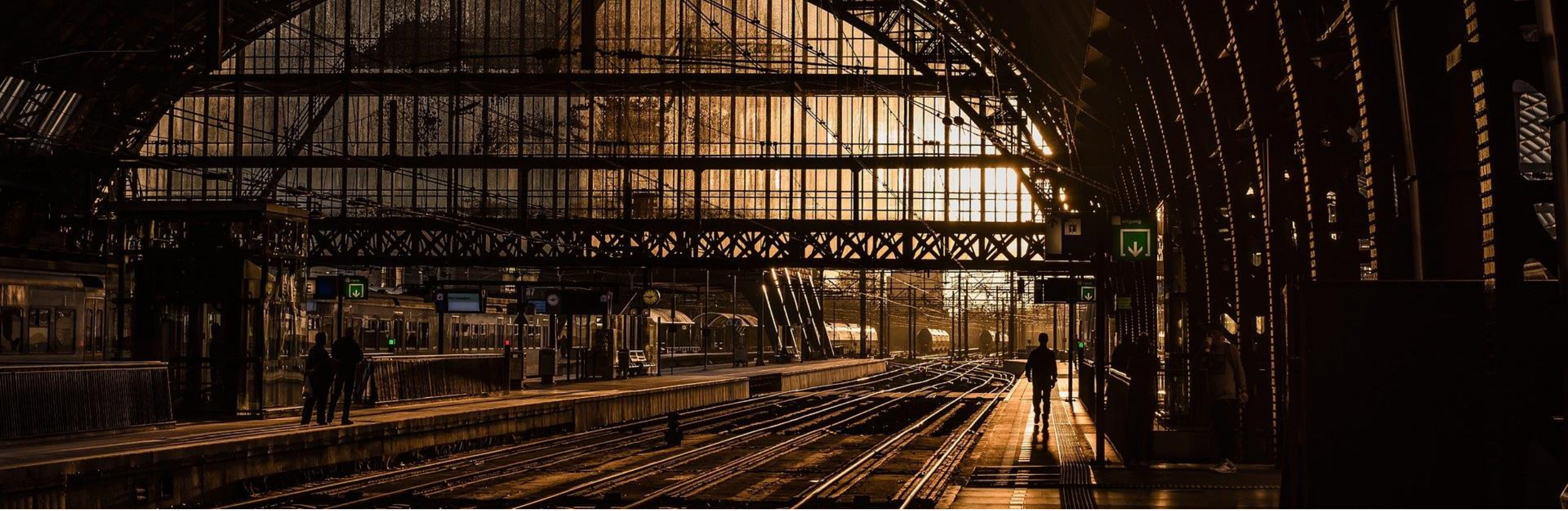


Concertation sur le décompte ferroviaire
Atelier de travail

23 novembre 2021



Bienvenue

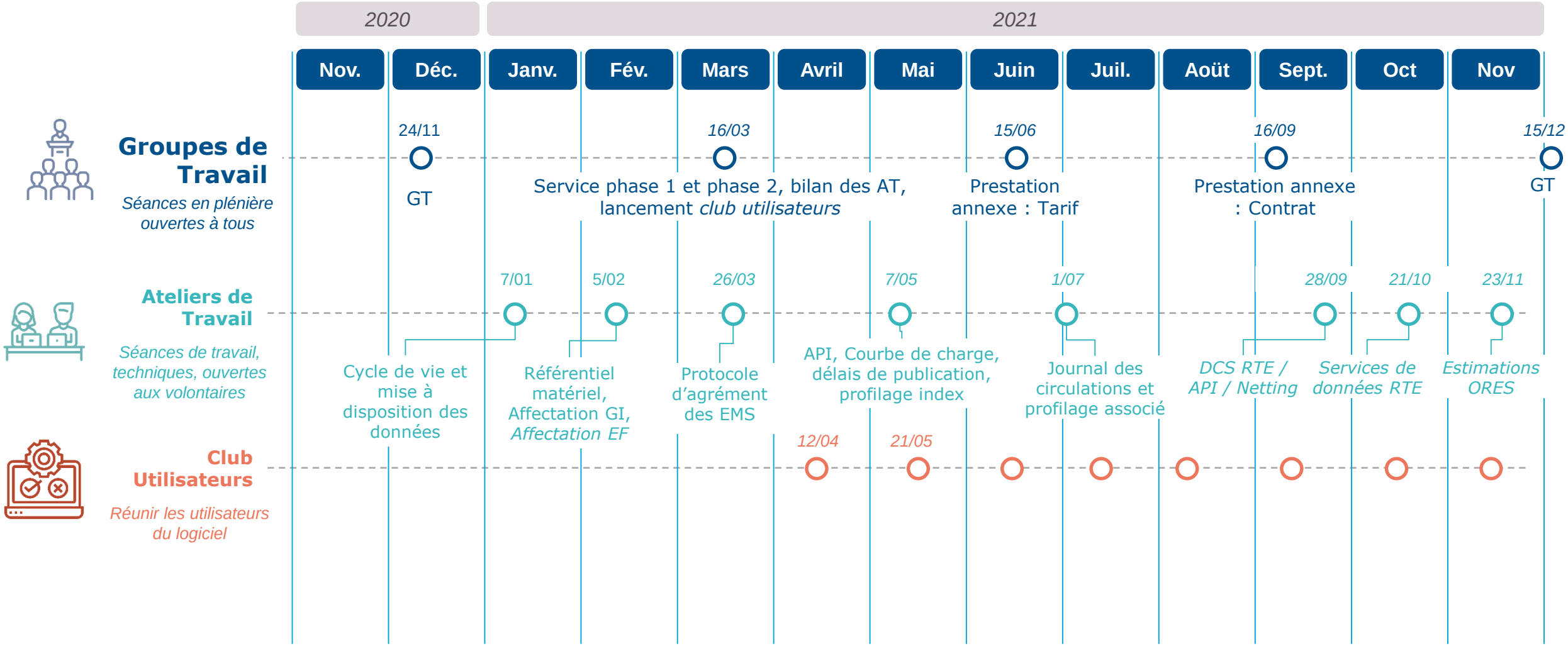
Sont présents aujourd'hui, sur invitation de RTE et SNCF Réseau :

- Captrain
- CNR
- EDF
- Euro Cargo Rail
- Europorte
- LFP
- LISEA
- SNCF Energie
- SNCF Fret
- SNCF Voyageurs
- UTP
- Thalys

Atelier de travail - modalités :

- Notre objectif aujourd'hui est d'échanger avec vous
- Nous allons vous présenter certains éléments, et vous poser des questions, le but est de récolter vos retours et vos besoins en séance

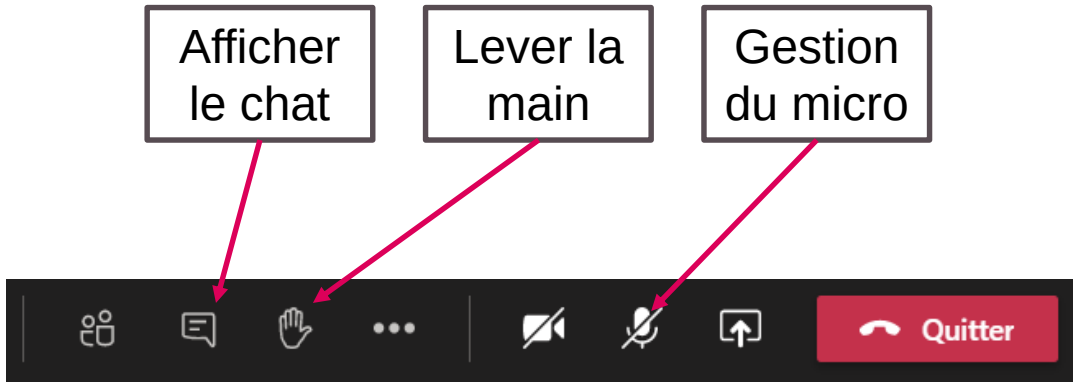
Planning concertation sur le décompte ferroviaire





Règles du jeu – atelier de travail

- Nous sommes nombreux, pour le confort de tous, merci de **penser à couper vos micros**.
- Pour rendre cet atelier **interactif**, nous vous invitons à **poser vos questions et à réagir au fur et à mesure**, par écrit (chat) ou en « levant la main ».
- Un **formulaire d'enquête** vous sera Excel vous sera partagé, que vous pourrez **renvoyer** après la présentation.





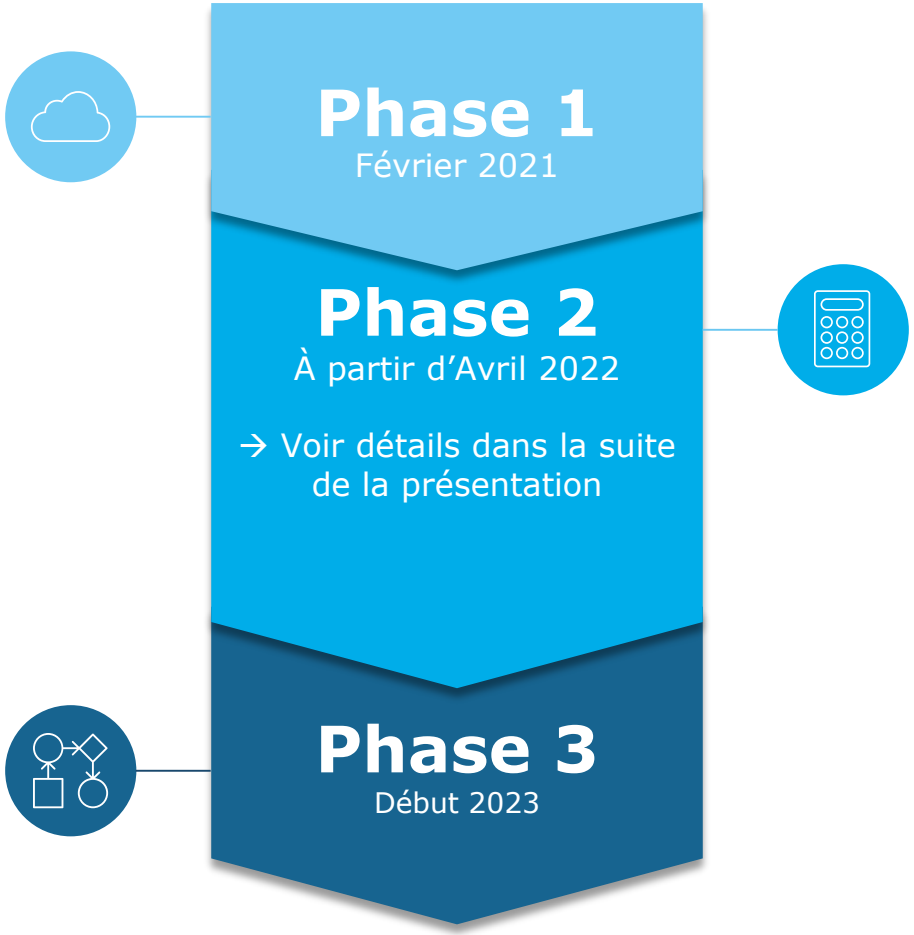
Agenda

1. **Introduction et rappels**
2. **Méthode d’estimation de la consommation**
3. **Critères de validation de l’algorithmie d’estimation**
4. **Processus de qualification, validation et de bascule**

Le projet de décompte ferroviaire SNCF R – RTE prévoit une transition progressive des rôles et responsabilités des activités de comptage et de décompte

1. **Nouveau** - Mise en service de l'IHM unique DECOFER (RTE) en remplacement de l'interface SOCLE (SNCF Réseau)
2. **Nouveau** – Mise en œuvre du DCS DECOFER. RTE est en capacité de télé-relever les dispositifs de comptage à la nouvelle version de la norme EN 50 463:2017.
3. SNCF Réseau reste responsable du calcul des décomptes et continue d'intégrer l'ensemble des données de comptage et réalise le calcul des décomptes, via l'outil Socle.

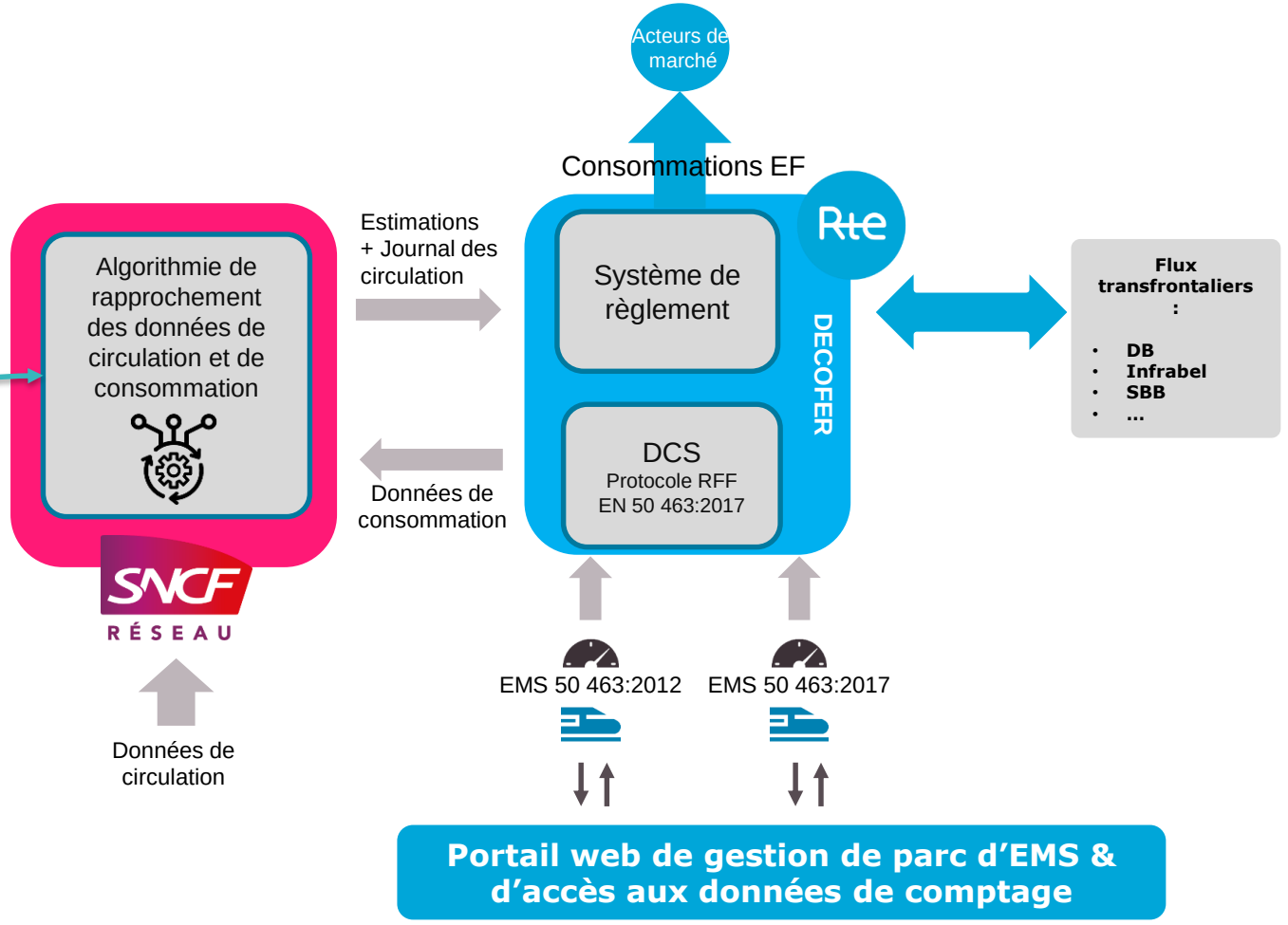
Nouveau – Les Gestionnaires d'Infrastructure deviennent les termes bouclants



1. **Nouveau** – RTE devient responsable du décompte et intègre, en remplacement de SNCF Réseau et via l'outil DECOFER, l'ensemble des données de comptage dans le calcul des décomptes.
2. **Nouveau** - SNCF Réseau assure l'estimation des engins non équipées de compteur, sur la base des données de circulation.
3. **Nouveau** – RTE devient l'opérateur de la télé-relève des compteurs anciens protocoles

Rappel et mise en contexte

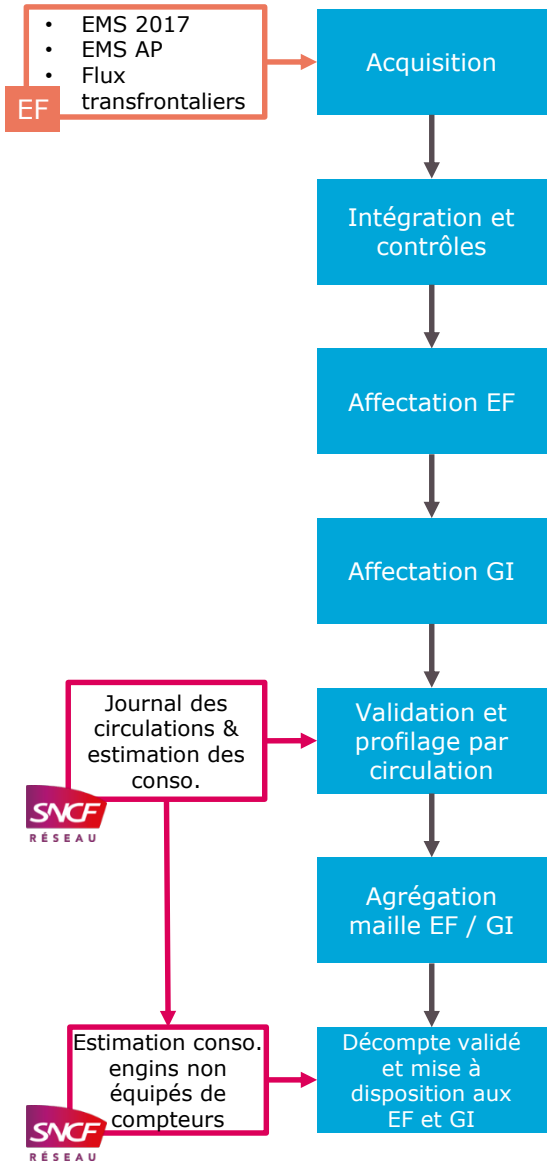
Schéma fonctionnel cible du décompte ferroviaire



Focus de l'atelier du 23/11

- Algorithmie de construction de l'estimation de la consommation

Rappels : rôle de l'estimation dans le calcul du décompte de la consommation des EF



L'estimation de la consommation des engins non équipés de chaque EF permet de compléter le décompte réalisé sur la base des engins équipés.

Cependant, les estimations n'ont pas vocation à remplacer les compteurs. L'équipement en compteurs reste une obligation réglementaire des EF, et un prérequis pour la fiabilisation du décompte.

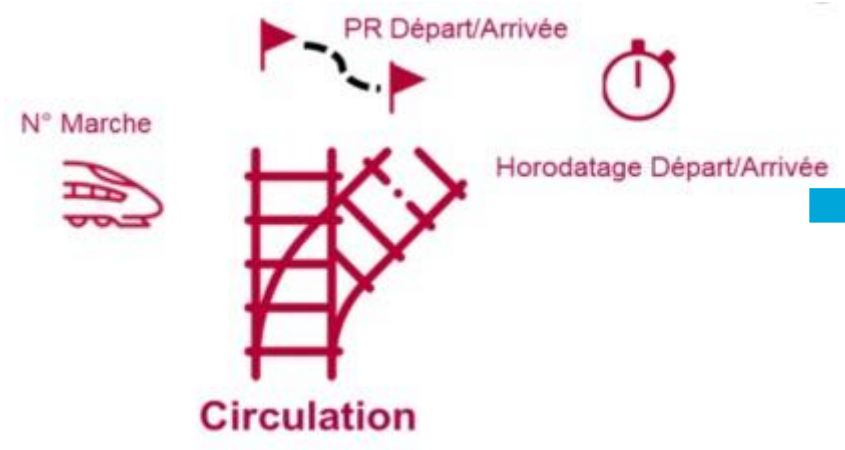
A ce stade le décompte obtenu concerne le parc d'**engins équipés de compteurs**

A ce stade le décompte final intègre **tous les engins**, en ajoutant **l'estimation** de la consommation pour les **engins non équipés** de compteurs

Rappel sur le journal des circulations (AT du 1/07)

Définition

Le **journal des circulations** est une liste des **circulations réalisées**, affectées à un **identifiant de l'engin moteur** (rame ou loco) et complétée d'une **estimation de la consommation** réalisée sur la circulation.



Rapprochement et identification *de l'engin moteur ayant réalisé la circulation*

L'engin moteur *est-il équipé d'un* ***compteur ?***

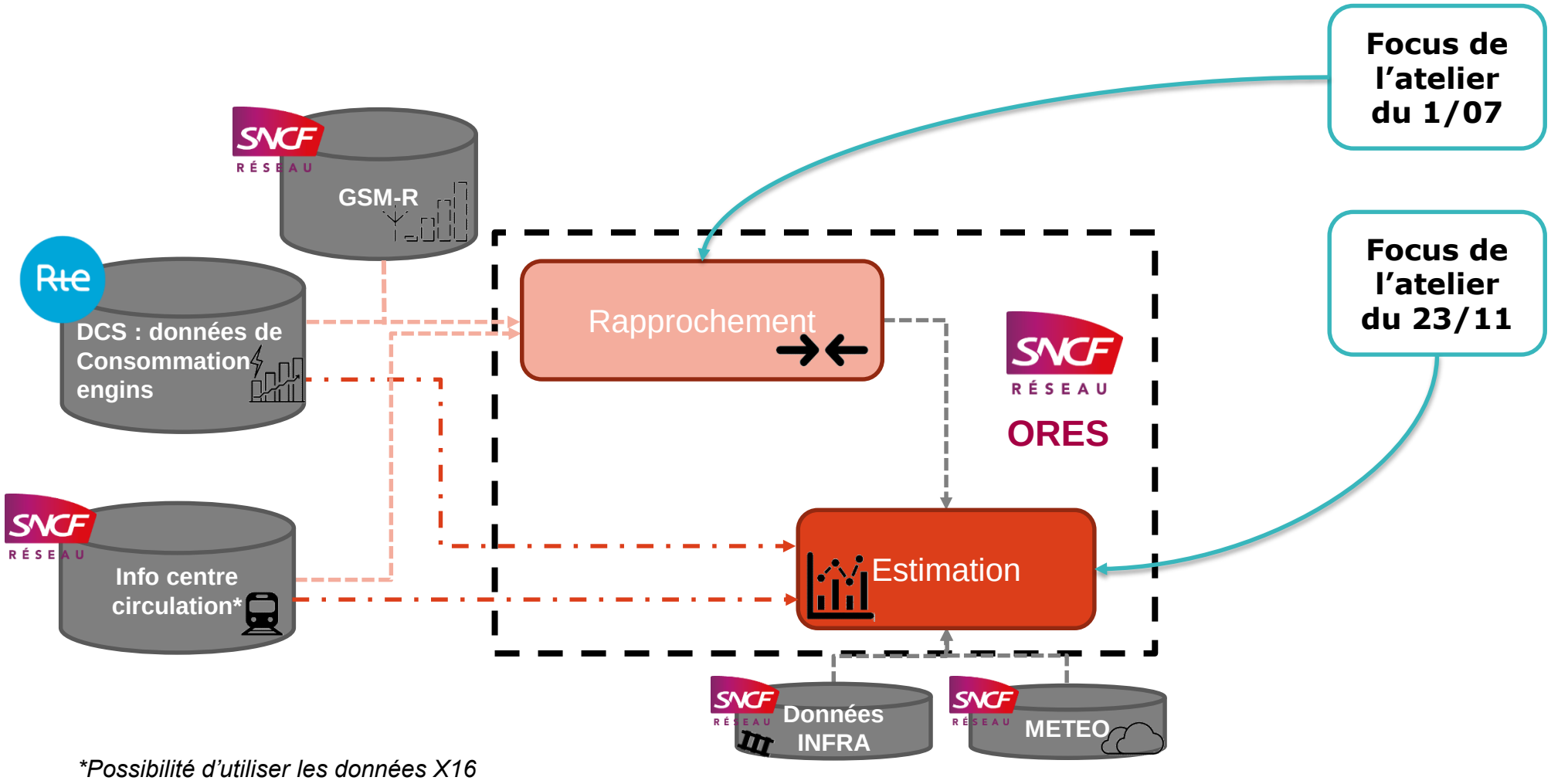
Estimation ***de la consommation*** *réalisée sur la circulation*

Numéro Marche	PR Départ	PR Arrivée	Horodatage Départ	Horodatage Arrivée	Identifiant engin moteur	Engin équipé de compteur ? (OUI / NON)	Estimation consommation
137743	Paris-St-Lazare	Argenteuil	01/02/2018 00:01	01/02/2018 00:15	93XXX09	NON	XX kWh
8419	Paris-Montparnasse	Bordeaux-St-Jean	01/02/2018 19:52	01/02/2018 21:51	92YYY25	OUI	YY kWh

Etape 1 ***Etape 2*** ***Etape 3***

Lors de l'atelier du 1er juillet 2021 nous avions présenté les étapes 1 (rapprochement), 2, et évoqué la 3. Aujourd'hui, nous revenons sur l'étape 3 de façon plus détaillée.

Rapprochement et estimation dans l'outil ORES



*Possibilité d'utiliser les données X16



Agenda

1. Introduction et rappels
2. Méthode d’estimation de la consommation
3. Critères de validation de l’algorithmie d’estimation
4. Processus de qualification, validation et de bascule

Construction de l'algorithme ORES

Les équipes de Data Scientist SNCF Réseau travaillent sur les algorithmes de l'outil ORES, avec pour objectif d'estimer la consommation des circulations.

Les choix méthodologiques, la conception et la validation (statistiques, machine learning, etc) sont réalisés en co-construction avec la Direction de l'Ingénierie de la Traction Electrique, sous le pilotage de la Direction de la Stratégie.

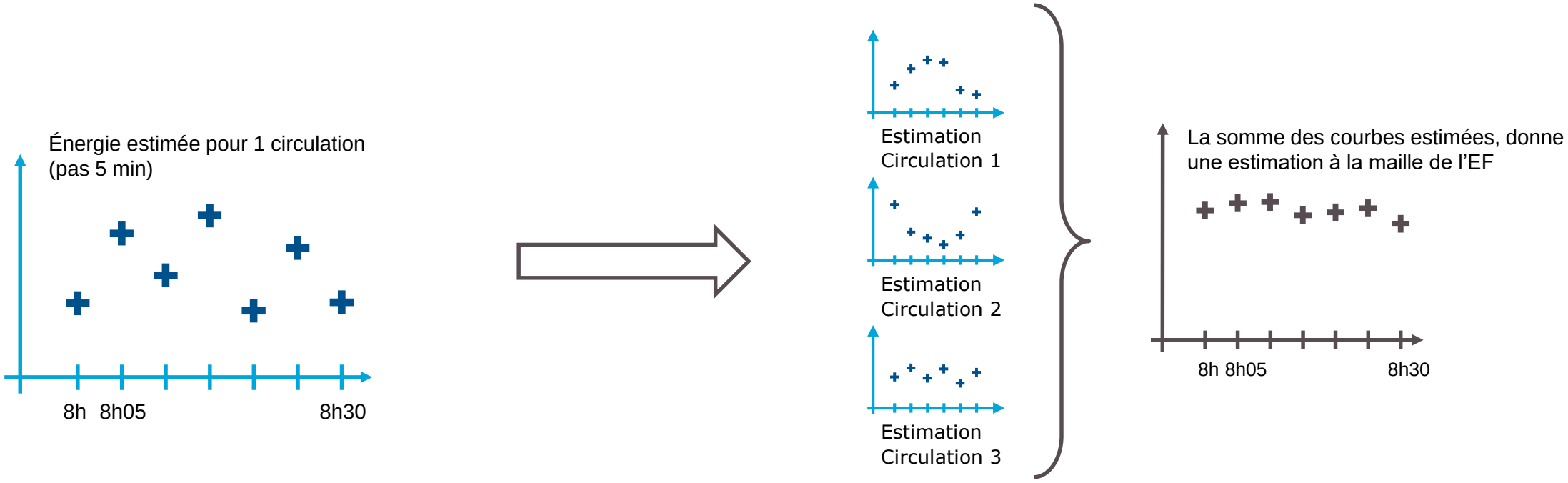
Les éléments présentés aujourd'hui constituent un état des travaux à date, l'outil ORES étant en cours d'optimisation.



Principes de construction de la courbe de charge des engins non équipés, à la maille d'une EF

Pour chaque circulation, l'outil ORES calcule une courbe de consommation estimée, au pas 5 min.

L'agrégation des courbes de consommation ainsi estimées permet d'obtenir l'estimation de la consommation de l'EF (à la maille de la partie non équipée de son parc)



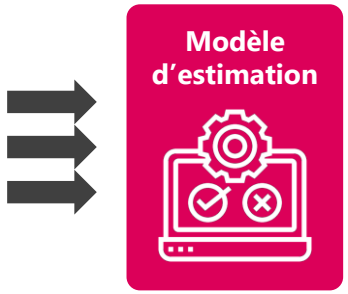
Dans la suite, on explique comment cette estimation, par circulation, au pas 5 min, est réalisée

Qu'est-ce qu'un modèle d'estimation ?

Un « modèle » d'estimation est une fonction permettant de calculer l'estimation de la consommation pour une circulation, sur la base d'un certain nombre de variables d'entrées.

Le modèle d'estimation fonctionne sur un pas de 5 min (pour les variables d'entrée et la conso calculée en sortie)

Données d'entrée pour une circulation non équipée de compteurs, au pas 5 min



Consommation estimée (au pas 5 min)

Les équipes SNCF Réseau ont construit à date 7 modèles, en se basant sur les « types de circulation » tels que définis dans le DRR

Trains régionaux, nationaux et internationaux de voyageurs aptes à la grande vitesse
Autres trains nationaux et internationaux de voyageurs
Trains régionaux de voyageurs (hors Transilien) non aptes à la grande vitesse
Trains régionaux de voyageurs Transilien non aptes à la grande vitesse
Trains fret
Autres trains (HLP, matériel…)
Tramway (*)

7 types de circulation → 7x



(*) Les tramways sont hors DRR mais ils ont des circulations et des caractéristiques un peu particuliers qui nécessitent donc un modèle d'estimation spécifique

Quelles variables d'entrée utiliser pour le modèle d'estimation ? (1/2)

Transfert de l'énergie électrique en énergie mécanique

Impact du type de conduite

$P_{électrique} = P_{traction} + P_{fonctionnement}$

Ce qui est mesuré par les compteurs

Terme lié aux consommations non liées à la traction : climatisation, chauffage, électronique…

La puissance de traction peut être <u>approchée</u> par la formule suivante :

$$P = \frac{1}{3,6}((A + Mgi' + kM\gamma)V + BV^2 + CV^3)$$

Avec :

Variable	Définition	Dépendance
A, B, C	Résistance à l'avancement	Type de train
M	Masse totale du train	Type de train, composition
k	Coefficient des masses tournantes	Type de train
i	Déclivité	Ligne circulée
γ	Accélération	Type de train, conducteur, TCT
V	Vitesse	Type de train, conducteur, TCT, ligne

Cette équation a permis d'identifier des variables pertinentes retenues comme données d'entrée dans les modèles d'estimation

Quelles variables d'entrée utiliser pour le modèle d'estimation ? (2/2)

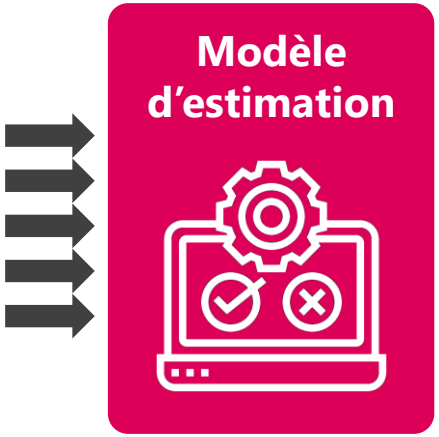
Parmi les variables listées précédemment, les variables suivantes, accessibles dans le SI Industriel SNCF Réseau, ont été retenues par les équipes SNCF Réseau.

Sources des données

- Infocentre
- GAIA
- Base Météo

Vitesse	Déclivité
Vitesse au carré	Virage (rayon de courbure)
Vitesse au cube	Nombre d'arrêt
Accélération	Vitesse minimale
Tonnage	Vitesse maximale
Température de l'air	

Données d'entrée pour une circulation non équipée de compteurs, au pas 5 min



On utilise le modèle d'estimation correspondant au type de circulation voulu

Consommation estimée (au pas 5 min)

Modèle d'estimation de la consommation : comment est « entrainé » le modèle ?

Entrainer un modèle permet de le paramétrer

Chaque modèle (1 par type de circulation) est entrainé sur la base
- des données d'entrée de circulation
- et des données de consommation mesurées par les compteurs.

Données d'entrée pour une circulation équipée de compteur, au pas 5 min

Vitesse	Déclivité
Vitesse au carré	Virage (rayon de courbure)
Vitesse au cube	Nombre d'arrêt
Accélération	Vitesse minimale
Tonnage	Vitesse maximale
Température de l'air	



*Le lien entre les **données d'entrée** (issues des circulations) et la **consommation mesurée** est possible grâce au **rapprochement**.*

Consommation mesurée par un compteur embarqué (pas 5 min)

Le modèle d'estimation est un modèle de type « arbre de décision », entraîné par la méthode « Random Forest »

Le bon fonctionnement des modèles d'estimation s'appuie sur deux grands pré-requis :

Les modèles d'estimations sont **entraînés sur les circulations** qui sont réalisées par des **engins équipés de compteurs**

→ Afin d'assurer la bonne représentativité de ces modèles, **les EF doivent être équipées de suffisamment de compteurs,** pour entrainer les modèles et les maintenir à jour



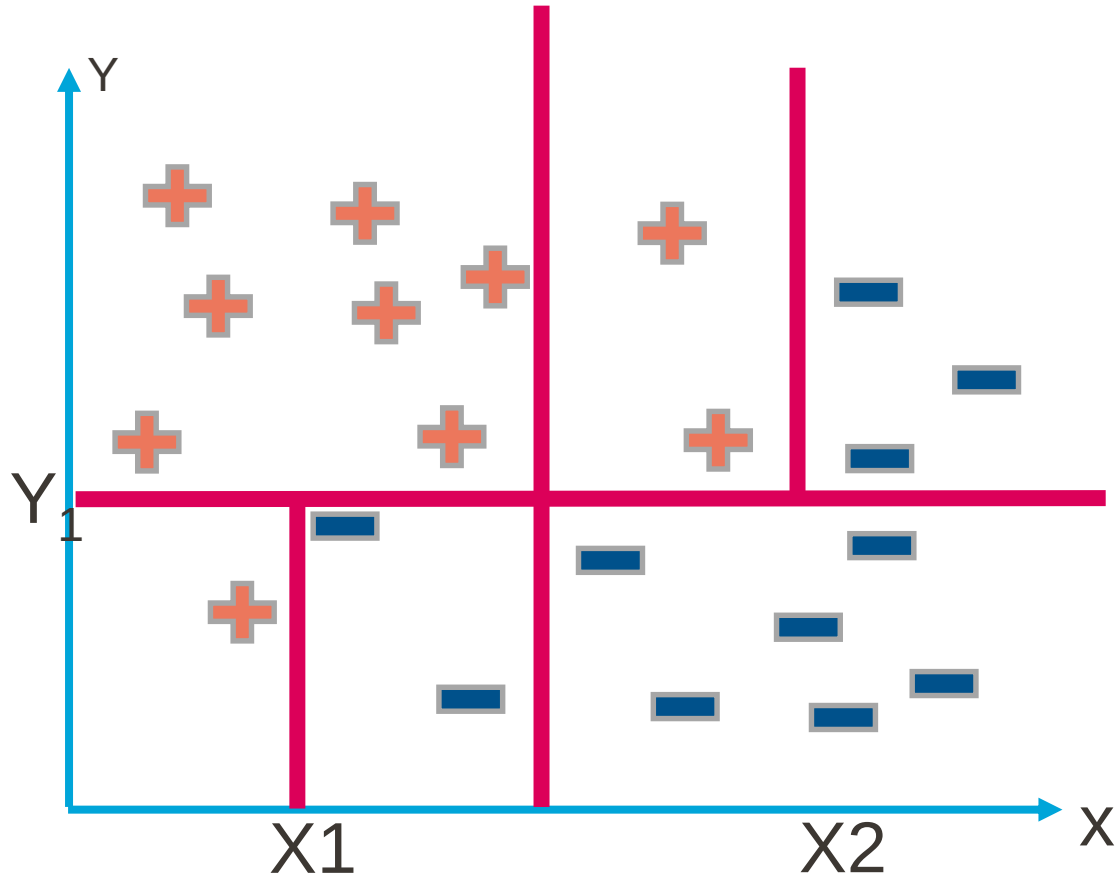


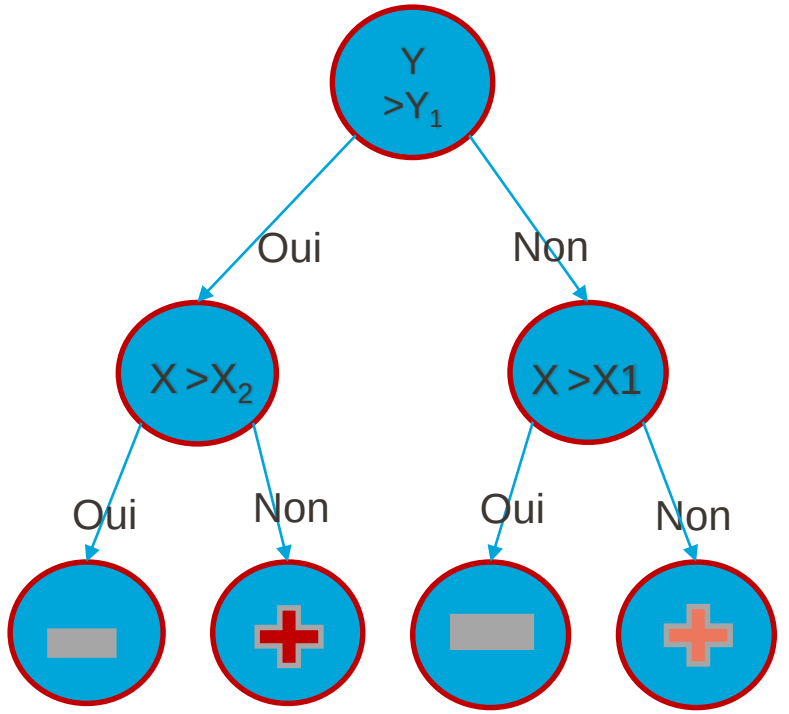
Déclarer avec précision le parc de matériel roulant (Référentiel matériel) est indispensable pour assurer le rapprochement et les estimations.

Par ailleurs, la précision des modèles d'estimations dépend également des **informations contenues dans les sillons** (tonnage, US/UM, type d'engin)

Le Modèle d'estimation de la consommation est de type « arbre de décision »

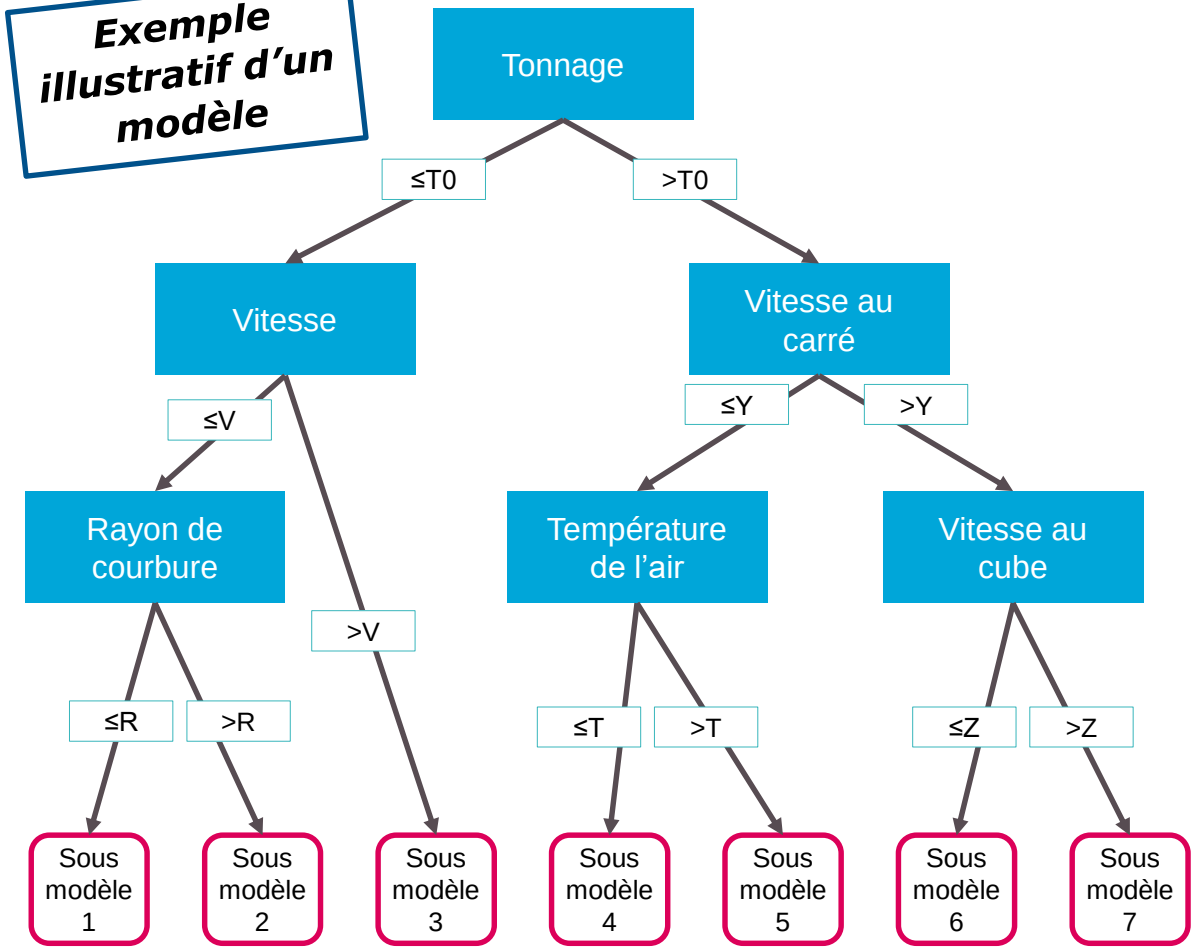
Schéma illustrant le principe d'un arbre de décision





Exemple de modèle de type « arbre de décision » obtenu

Exemple illustratif d'un modèle



L'arbre de décision permet d'aboutir à plusieurs « sous-modèles » : un par case

Pour réaliser l'estimation de la consommation d'un train sur un pas 5 min : sur un pas de temps, l'arbre de décision du modèle (lié au type de circulation) est parcouru, et les seuils définis à chaque nœud permettent ***d'aboutir au choix du bon « sous modèle »***

Le sous-modèle ***permet ensuite de calculer la consommation estimée****. Il est basé sur une fonction linéaire (type : y = a.x + b)*

Pour un type de circulation, le choix des variables de seuil, des seuils et des sous-modèles ***est fixé par entrainement du modèle.***



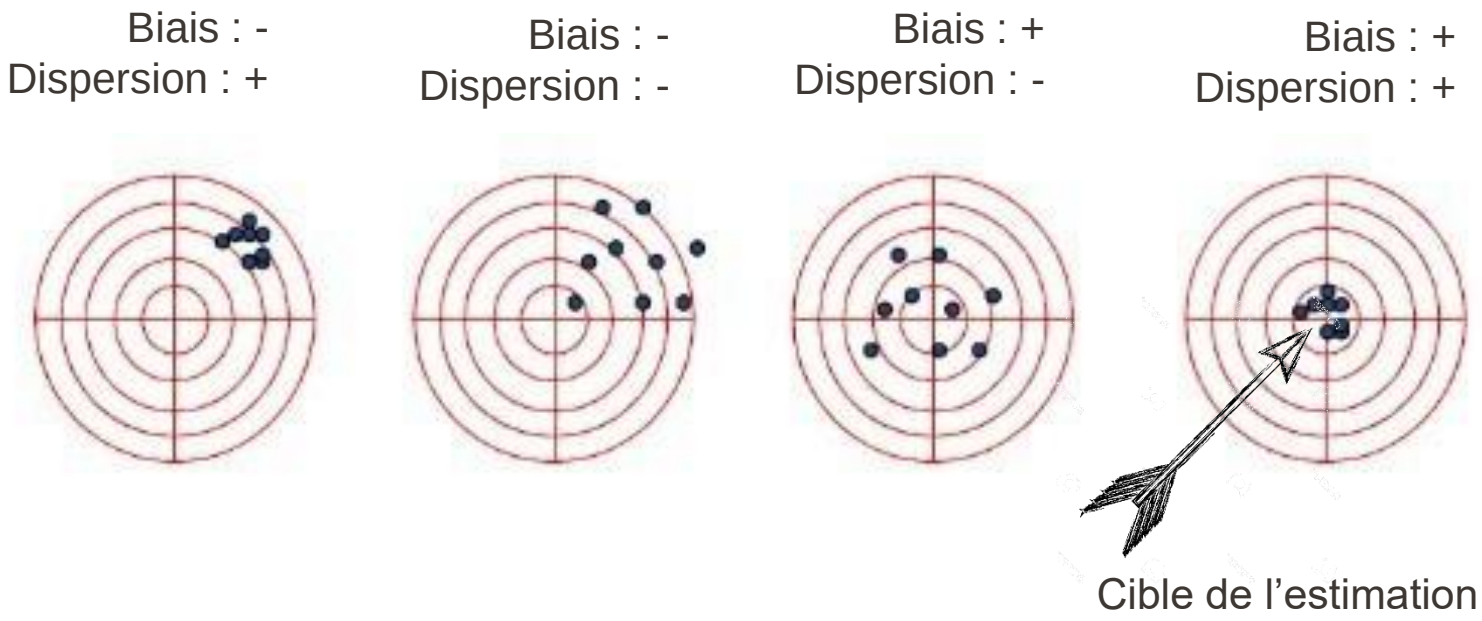
Agenda

1. **Introduction et rappels**
2. **Méthode d’estimation de la consommation**
3. **Critères de validation de l’algorithmie d’estimation**
4. **Processus de qualification, validation et de bascule**

Qu’est ce qu’une bonne estimation?

Les deux critères principaux: le biais et la dispersion

- Le biais est une évaluation de l’écart entre les valeurs d’énergie estimées et les valeurs réelle mesurées.
- La dispersion décrit la variabilité des valeurs d’estimation.



Les premières objectifs de validation

Evaluation du biais et de la dispersion

- Le biais peut s'estimer via la moyenne des valeurs par niveau d'agrégation.
- Les dispersions peuvent s'estimer en utilisant le rapport entre l'écart type et la médiane d'une estimation.

La validation sera réalise en agrégeant à différents niveaux

1. Sur une série de circulations identiques
2. Au périmètre d'un type de circulation
3. Au périmètre d'une EF, sur un mois.



Les sensibilités seront aussi vérifiées par rapport aux variables : Masse, Vitesse, temps de parcours

Les premières objectifs de validation

Vérification que l'étendue de variation des estimations correspondent à l'étendue de variation des mesures sur une série de circulations identiques

num_ circulation	DATE_HEURE_ DEPART	DATE_HEURE_ ARRIVEE	PR_ origine	PR_ destination	NUM_ENGIN_ Rapprochés	conso_ Estimé	conso_ Mesurée	Ecart (Estimée-Mesurée)	Ecart absolu
1234	10/08/2020 09:09	10/08/2020 11:08	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 613,61	2 570,60	1,67	1,67
1234	11/08/2020 09:08	11/08/2020 11:12	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 746,48	2 795,70	-1,76	1,76
1234	12/08/2020 09:07	12/08/2020 11:11	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 911,18	2 585,70	12,59	12,59
1234	13/08/2020 09:05	13/08/2020 11:10	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 932,05	2 301,80	27,38	27,38
1234	14/08/2020 09:05	14/08/2020 11:10	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 681,97	2 494,40	7,52	7,52
1234	18/08/2020 09:06	18/08/2020 11:09	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 569,10	2 575,90	-0,26	0,26
1234	19/08/2020 09:11	19/08/2020 11:10	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 974,68	2 675,90	11,17	11,17
1234	20/08/2020 09:05	20/08/2020 11:10	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 882,14	2 663,20	8,22	8,22
1234	21/08/2020 09:05	21/08/2020 11:12	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 979,69	2 294,70	29,85	29,85
1234	25/08/2020 09:04	25/08/2020 11:14	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 990,94	2 389,00	25,20	25,20
1234	26/08/2020 09:04	26/08/2020 11:07	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 618,25	2 829,60	-7,47	7,47
1234	27/08/2020 09:04	27/08/2020 11:09	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 655,37	2 402,50	10,53	10,53
1234	28/08/2020 09:16	28/08/2020 11:12	Ville 1	Ville 2	9x87xxxxxxxxx_9x87xxxxxxxxx	2 472,81	2 558,30	-3,34	3,34
Moyenne totale						**2 771,41**	**2 549,02**	**9,33**	**11,30**

Mesurée	max	2 829,60
	min	2 294,70
	Intervalle de variation cumulé(%)	**10,40**

Estimée	max	2 990,94
	min	2 472,81
	Intervalle de variation cumulé(%)	**9,43**

Biais en %

Dispersion

Critères de validation:

- L'estimation est cohérente avec la consommation mesurée. Elle varie suivant la variation de cette dernière.
- Sur ces 13 exemples du même trajet, la conso mesurée varie de ±10,40% autour de la médiane.
- La variation de l'estimation de ±9,43% est située dans cet intervalle.



Agenda

1. **Introduction et rappels**
2. **Méthode d'estimation de la consommation**
3. **Critères de validation de l'algorithmie d'estimation**
4. **Processus de qualification, validation et de bascule**

Quelles étapes avant la mise en œuvre opérationnelle des estimations ORES dans le décompte ?

Janvier 2022 : Mise en Production DECOFER & ORES interconnectés

- Janvier 2022 – avril 2022 : **Observation** et évaluation de la performance des modèles d'estimation sur la base des données réelles
  - Les résultats de l'observation seront partagés avant le démarrage de la marche en double.

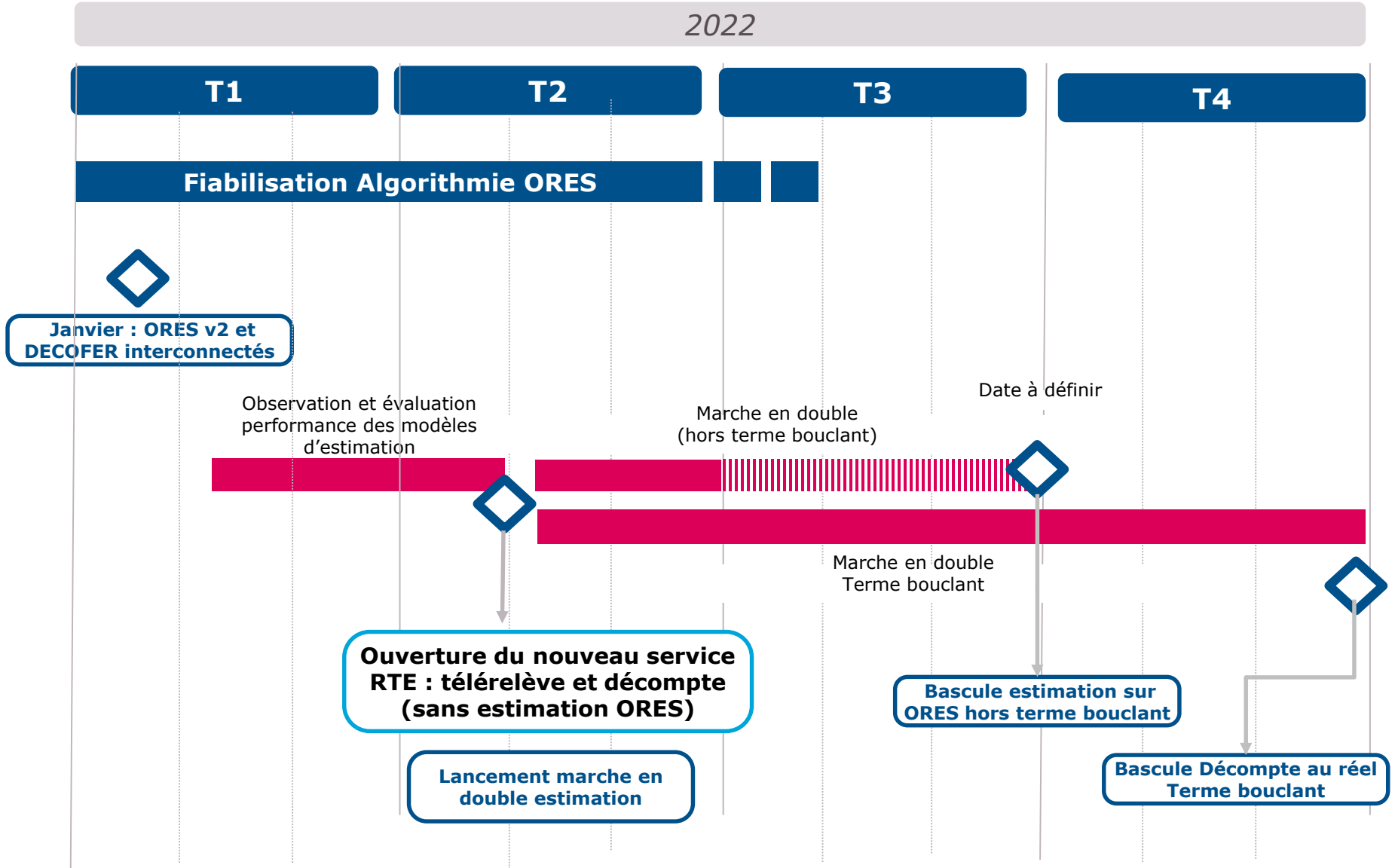
Avril 2022 - Mise en place de la marche en double

- Dans cette phase (durée : minimum 3 mois), le décompte officiel publié aux acteurs de marché reste le décompte sans données ORES.
  - En parallèle, RTE et SNCF Réseau mettent à disposition des acteurs qui le souhaitent, le futur décompte, incluant les données ORES

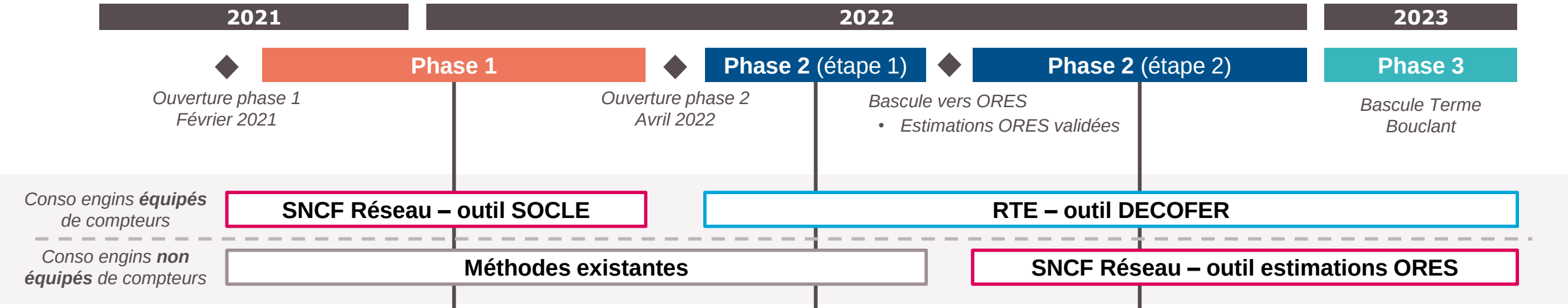
- La marche en double permettra
  - De confirmer les modèles
  - D'évaluer l'impact du nouveau calcul de la consommation de chaque EF, par rapport au calcul existant
  - D'analyser les éventuels écarts

Date à déterminer (S2 2022) – le décompte inclut ORES

Planning de mise en œuvre du nouveau service



Évolutions des périmètres de services en 2022



Phase 1 du service depuis le 22/02/2021

SNCF Réseau reste responsable du décompte ferroviaire

- SNCF Réseau continue de télérelever les EMS anciens protocoles
- SNCF Réseau continue de réaliser le calcul des décomptes
- RTE télé-relève les EMS 2017
- RTE met à disposition le nouveau portail Web de gestion des Engins et EMS

Phase 2 – étape 1 à partir d'avril 2022

RTE devient responsable du décompte ferroviaire

- RTE Télérelève les EMS 2017 et les protocoles RFF
- RTE calcule la consommation des engins équipés de compteurs

Phase 2 – étape 2

RTE reste responsable du décompte ferroviaire
SNCF Réseau assure l'estimation des engins non équipés

- SNCF Réseau fournit le journal des circulations et l'estimation de la consommation des engins non équipés de compteurs, calculés avec ORES
- RTE intègre cette estimation dans la consommation de l'EF en l'ajoutant à celle des engins équipés



Merci de votre attention