

19 RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
COURBEVOIE

11 N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

3 083 355

21 N° d'enregistrement national : 18 56007

51 Int Cl⁸ : G 06 Q 10/04 (2018.01), G 06 F 17/10, G 16 Z 99/00,
G 06 Q 50/06

12 DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 29.06.18.

30 Priorité :

43 Date de mise à la disposition du public de la
demande : 03.01.20 Bulletin 20/01.

56 Liste des documents cités dans le rapport de
recherche préliminaire : *Ce dernier n'a pas été
établi à la date de publication de la demande.*

60 Références à d'autres documents nationaux
apparentés :

○ Demande(s) d'extension :

71 Demandeur(s) : EQUENSWORLDFINE S.E. — NL.

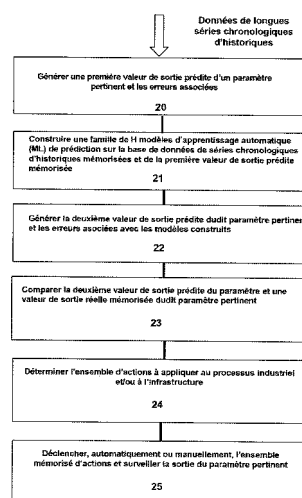
72 Inventeur(s) : HATTAB Dalila.

73 Titulaire(s) : EQUENSWORLDFINE S.E..

74 Mandataire(s) : DEBAY IP.

54 PROCÉDE, SYSTEME ET PRODUIT POUR L'AMELIORATION D'UN PROCESSUS INDUSTRIEL.

57 L'invention concerne un procédé, un système et un programme d'ordinateur pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle, le système comprenant au moins un ensemble de capteurs pour capturer au moins des valeurs de sortie d'au moins un paramètre pertinent du processus industriel, une base de données pour mémoriser lesdites valeurs de sortie capturées et au moins des données de longues séries chronologiques d'historique relatives au paramètre pertinent ou au processus industriel, au moins une infrastructure informatique comprenant au moins un processeur sur lequel sont exécutés un ensemble de programmes mémorisés dans une mémoire pour mettre en oeuvre ledit procédé comprenant au moins la prédiction et la mémorisation d'au moins une valeur de sortie d'un paramètre pertinent et le déclenchement d'un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle ledit processus industriel s'exécute.



FR 3 083 355 - A1



**Procédé, système et produit pour l'amélioration d'un processus
industriel**

DOMAINE TECHNIQUE DE L'INVENTION

5 La présente invention concerne le domaine de la surveillance de processus industriels et, plus précisément, un procédé, un système et un produit informatique pour améliorer le rendement desdits processus industriels. Ces procédé, système et produit informatique pourraient être appliqués à la surveillance d'un processus industriel se rapportant, par
10 exemple, et de manière non limitative, à l'hydrologie et particulièrement à l'hydroélectricité.

ARRIERE-PLAN TECHNOLOGIQUE DE L'INVENTION

 La prédiction d'une valeur de sortie future d'un processus physique tel
15 que la génération d'énergie électrique par un dispositif ou une machine tel qu'un réacteur nucléaire, par exemple, est une clé pour l'élaboration d'un ensemble de stratégies en vue d'améliorer le rendement ou le taux de production du dispositif ou de la machine. Par exemple, la prédiction de l'intensité de la vitesse du vent peut être utile pour la gestion des risques
20 climatiques.

 Il existe, dans l'art antérieur, des procédés et un système pour prévoir les valeurs futures d'un paramètre physique donné, cependant, lesdits procédés sont souvent limités par leur capacité à ne prédire qu'un phénomène avec des données d'historique ou de mémoire courte. En effet, une prédiction
25 de données d'historique ou de mémoire longue est souvent nécessaire pour capturer la variabilité ou la volatilité d'un paramètre physique pertinent pour améliorer un processus industriel.

DESCRIPTION GENERALE DE L'INVENTION

La présente invention a pour but de pallier certains inconvénients de l'art antérieur en proposant une procédure en vue d'augmenter au moins la productivité d'un processus industriel.

- 5 Ce but est atteint par un procédé pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle comprenant au moins une base de données comprenant des données de longues séries chronologiques d'historique du processus industriel ou d'au moins un paramètre pertinent dudit processus industriel, au moins une
- 10 infrastructure informatique contenant au moins une mémoire pour mémoriser un ensemble de programmes, au moins un processeur pour exécuter ledit procédé comprenant au moins une étape consistant à :
- générer des premières valeurs de sortie prédites d'un paramètre pertinent du processus industriel avec un premier modèle de
 - 15 prédiction mémorisé et les erreurs associées, sur la base des données de longues séries chronologiques d'historique relatives audit paramètre pertinent mémorisées dans la base de données,
- le procédé étant caractérisé en ce qu'il comprend également au moins les étapes suivantes consistant à :
- 20 • construire une famille de H modèles d'apprentissage automatique (ML) de prédiction, constituant le deuxième modèle de prédiction, pour la prédiction de multiples valeurs de sortie du même paramètre pertinent sur la base des données de séries chronologiques d'historique mémorisées et des premières valeurs
 - 25 de sortie prédites mémorisées et générer, avec ledit deuxième modèle de prédiction, de multiples deuxièmes valeurs de sortie prédites dudit paramètre pertinent et les erreurs associées ;
 - comparer les multiples deuxièmes valeurs de sortie prédites du paramètre pertinent et les valeurs de sortie réelles mémorisées
 - 30 dudit paramètre pertinent ;

- déterminer l'ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute, sur la base de l'analyse de la comparaison, et mémoriser lesdites actions ;
- 5
 - déclencher, automatiquement ou manuellement, l'ensemble d'actions mémorisé et surveiller la sortie du paramètre pertinent sur la base au moins des erreurs prédites obtenues à partir de la deuxième étape de prédiction.
- 10

Selon une autre particularité, le premier modèle de prédiction utilisé pour générer des premières valeurs de sortie prédites d'un paramètre pertinent est un modèle ARFIMA (modèle autorégressif à moyennes mobiles fractionnairement intégré).
- 15

Selon une autre particularité, chaque modèle de la famille de H modèles d'apprentissage automatique de prédiction utilisé pour prévoir les multiples deuxièmes valeurs de sortie prédites du paramètre pertinent est construit au moyen d'une approche de stratégie directe élargie dans laquelle les premières valeurs de sortie prédites mémorisées sont incorporées en tant que variables
- 20

d'entrée exogènes en plus des données de longues séries chronologiques d'historique mémorisées relatives au paramètre pertinent du processus industriel.
- Selon une autre particularité, le modèle d'apprentissage automatique
- 25

est un modèle ANN (réseau neuronal artificiel).
- Selon une autre particularité, le modèle d'apprentissage automatique
- est un modèle k-NN (k voisins les plus proches).
- 30

Selon une autre particularité, le modèle d'apprentissage automatique est un modèle SVR (régression basée sur une machine à vecteur de support).

Selon une autre particularité, le nombre de jour minimum, D, des données d'historique mémorisées pour améliorer le rendement du processus industriel est compris dans un intervalle défini par $100 < D < 300$.

- 5 Un autre but de la présente invention consiste à proposer une infrastructure pour augmenter au moins la productivité d'un processus industriel.

10 Cet but est atteint par un système pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle, le système comprenant au moins un ensemble de capteurs pour capturer au moins des valeurs de sortie d'au moins un paramètre pertinent du processus industriel, une base de données pour mémoriser lesdites données de sortie capturées et au moins des données de longues séries chronologiques
15 d'historique relatives au paramètre pertinent ou au processus industriel, au moins une infrastructure informatique comprenant au moins un processeur sur lequel sont exécutés un ensemble de programmes mémorisés dans une mémoire pour prédire et mémoriser au moins de multiples valeurs de sortie d'un paramètre pertinent et déclencher, automatiquement ou manuellement,
20 un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle ledit processus industriel s'exécute, au moyen de la mise en œuvre du procédé pour améliorer le rendement d'un processus industriel tel que décrit dans la présente demande.

- 25 Selon une autre particularité, le système comprend un dispositif de prédiction contenant au moins :

- une première unité de prédiction comprenant au moins une première mémoire pour mémoriser des données, un processeur exécutant le premier modèle de prédiction pour calculer les
30 premières valeurs de sortie prédites du paramètre pertinent et les mémoriser dans la première mémoire ;

- 5 • une unité de modélisation comprenant au moins une deuxième mémoire pour mémoriser des données, un processeur pour exécuter un ensemble de programmes pour construire le deuxième modèle de prédiction de H modèles d'apprentissage automatique et les mémoriser dans la deuxième mémoire ;
- 10 • une deuxième unité de prédiction comprenant une troisième mémoire pour mémoriser des données, un processeur pour exécuter le deuxième modèle de prédiction pour calculer les multiples deuxième valeurs de sortie prédites du paramètre pertinent et les erreurs prédites associées, et mémoriser lesdites valeurs de sortie et les erreurs prédites dans la troisième mémoire.

15 Selon une autre particularité, le système comprend un dispositif d'analyse de données comprenant au moins un processeur, au moins une mémoire pour mémoriser des données et/ou un ensemble de programmes dont l'exécution sur le processeur met en œuvre la comparaison des multiples deuxième valeurs de sortie prédites du paramètre pertinent et des valeurs de sortie réelles dudit paramètre pertinent et génère au moins un ensemble
 20 d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

25 Selon une autre particularité, le dispositif d'analyse de données transmet le résultat de l'analyse à une interface interactive d'un opérateur du processus industriel, ladite interface interactive comprenant au moins un agencement pour permettre à l'opérateur de générer, après avoir vérifié le résultat de l'analyse, un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

30 Selon une autre particularité, le système comprend un dispositif de déclenchement, connecté à au moins une unité de commande du processus industriel, contenant au moins un processeur, au moins une mémoire pour

mémoriser des données et/ou des instructions exécutables sur le processeur de manière à déclencher l'ensemble d'actions mémorisé à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

5

Selon une autre particularité, le dispositif de déclenchement comprend un agencement pour appliquer automatiquement l'ensemble d'actions au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

10

Selon une autre particularité, le dispositif de déclenchement comprend un agencement pour appliquer manuellement l'ensemble d'actions au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

15

Selon une autre particularité, le dispositif de déclenchement comprend également une alarme pour alerter un opérateur du fait qu'un ensemble d'actions doivent être appliquées audit processus.

20 Selon une autre particularité, la première unité de prédiction est connectée, par l'intermédiaire de moyens de communication, à l'unité de modélisation de manière à transmettre au moins les premières valeurs de sortie prédites du paramètre pertinent du processus industriel, lesdites valeurs de sortie prédites étant mémorisées dans la deuxième mémoire.

25 Selon une autre particularité, l'unité de modélisation est connectée, par l'intermédiaire de moyens de communication, à la deuxième unité de prédiction de manière à transmettre le deuxième modèle de prédiction de H modèles d'apprentissage automatique, ledit deuxième modèle de prédiction étant mémorisé dans la troisième mémoire.

30

Selon une autre particularité, la première unité de prédiction est connectée, par l'intermédiaire de moyens de connexion, à la deuxième unité

de prédiction de manière à transmettre les premières valeurs de sortie prédites du paramètre pertinent du processus industriel utilisées en tant que variables d'entrée pour le deuxième modèle de prédiction, lesdites valeurs de sortie prédites étant mémorisées dans la troisième mémoire.

5

Selon une autre particularité, la deuxième unité de prédiction est connectée, par l'intermédiaire de moyens de communication, au dispositif d'analyse de données de manière à transmettre au moins les multiples deuxièmes valeurs de sortie prédites du paramètre pertinent, et les erreurs
10 prédites relatives auxdites multiples deuxièmes valeurs de sortie prédites.

Selon une autre particularité, le dispositif d'analyse de données est connecté, par l'intermédiaire de moyens de communication, au dispositif de déclenchement de manière à transmettre un résultat de l'analyse, et
15 l'ensemble d'actions à appliquer.

Selon une autre particularité, la première unité de prédiction, la deuxième unité de prédiction, l'unité de modélisation et au moins le dispositif d'analyse de données sont connectés à la base de données de manière à
20 récupérer les données d'historique et/ou les données réelles relatives au processus industriel ou au paramètre pertinent dudit processus industriel.

Selon une autre particularité, les moyens de communication de la première unité de prédiction, de la deuxième unité de prédiction, de l'unité de
25 modélisation, du dispositif d'analyse de données et au moins du dispositif de déclenchement sont des moyens de communication sans fil/filaires.

Un autre but de la présente invention consiste à proposer au moins un ensemble d'instructions qui, lorsqu'elles sont exécutées sur au moins un
30 processeur, permettent d'augmenter au moins la productivité d'un processus industriel.

Cet but est atteint par un produit-programme d'ordinateur, mémorisé sur un support pouvant être lu par des moyens informatiques et contenant des instructions exécutables par au moins un processeur de ces moyens informatiques pour mettre en œuvre le procédé tel que décrit dans la présente
5 demande.

DESCRIPTION DES FIGURES ILLUSTRATIVES

D'autres caractéristiques et avantages de la présente invention apparaîtront plus clairement à la lecture de la description ci-après, faite en
10 référence aux annexés, dans lesquels :

- la figure 1 est une représentation schématique des étapes du procédé pour améliorer le rendement du processus industriel, selon un mode de réalisation ;

- la figure 2 est une représentation schématique du système pour
15 améliorer le rendement du processus industriel, selon un mode de réalisation ;

- la figure 3 est une représentation schématique de tableaux comparant les approches de stratégie directe (DirStr) et de stratégie directe élargie (DirStrX) ;

- les figures 4a et 4b sont respectivement des représentations
20 schématiques des vrais (réels) rendements de l'indice CAC40 sur la période couvrant la crise de 2002 jusqu'à 2010 et du tableau des mesures d'erreurs et des performances de trading pour tous les procédés ;

- les figures 5a et 5b sont respectivement des représentations
25 schématiques des vrais (réels) rendements de l'indice S&P 500 sur la période couvrant la crise de 2002 jusqu'à 2010, et du tableau des mesures d'erreurs et des performances de trading pour tous les procédés ;

- les figures 6a et 6b sont respectivement des représentations
30 schématiques des vrais (réels) rendements de l'indice CAC40 sur la période post-crise de 2015 jusqu'à 2017 et du tableau des mesures d'erreurs et des performances de trading pour tous les procédés ;

- les figures 7a et 7b sont respectivement des représentations schématiques des vrais (réels) rendements de l'indice S&P 500 sur la période post-crise de 2015 jusqu'à 2017, et du tableau des mesures d'erreurs et des performances de trading pour tous les procédés.

5

DESCRIPTION DES MODES DE REALISATION PREFERES DE L'INVENTION

La présente invention concerne un procédé pour améliorer le rendement d'un processus industriel.

10

Dans certains modes de réalisation, le procédé (figure 1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle comprenant au moins une base de données comprenant des données de longues séries chronologiques d'historique du processus industriel ou d'au moins un paramètre pertinent dudit processus industriel, au moins une infrastructure informatique contenant au moins une mémoire pour mémoriser un ensemble de programmes, au moins un processeur pour exécuter ledit procédé comprenant au moins une étape consistant à :

15

20

- générer des premières valeurs de sortie prédites d'un paramètre pertinent du processus industriel avec un premier modèle de prédiction mémorisé et les erreurs associées, sur la base des données de longues séries chronologiques d'historique relatives audit paramètre pertinent mémorisées dans la base de données ;

25

le procédé étant caractérisé en ce qu'il comprend également au moins les étapes suivantes consistant à :

30

- construire une famille de H modèles d'apprentissage automatique (ML) de prédiction, constituant le deuxième modèle de prédiction, pour la prédiction de multiples valeurs de sortie du même paramètre pertinent sur la base des données de séries chronologiques mémorisées et des premières valeurs de sortie

prédites mémorisées et générer, avec ledit deuxième modèle de prédiction, de multiples deuxièmes valeurs de sortie prédites dudit paramètre pertinent et les erreurs associées ;

- 5
 - comparer les multiples deuxièmes valeurs de sortie prédites du paramètre pertinent et les valeurs de sortie réelles mémorisées dudit paramètre pertinent ;
 - déterminer l'ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute, sur la base de l'analyse de la comparaison
- 10
 - et mémoriser lesdites actions ;
 - déclencher, automatiquement ou manuellement, l'ensemble d'actions mémorisé et surveiller la sortie du paramètre pertinent sur la base au moins des erreurs prédites obtenues à partir de la deuxième étape de prédiction.

15

Par exemple, et de manière non limitative, le processus industriel peut être lié, dans certains modes de réalisation, au domaine de l'hydrologie. En effet, le procédé proposé peut être utilisé pour prévoir (prédire) l'approvisionnement en eau et les demandes d'eau, des jours, semaines, mois

20 et années, en lien avec les données d'historiques hydrométéorologiques à mémoire longue ou mémorisées. Lesdites données d'historique hydrométéorologiques peuvent comprendre, par exemple, la pluviométrie, le débit, la température, l'humidité, l'équivalent en eau de la neige, l'épaisseur de neige, la couverture nuageuse, le rayonnement, l'évaporation ou

25 l'évapotranspiration, les niveaux d'eau dans le sol et le vent. L'infrastructure mettant en œuvre le processus industriel d'alimentation (approvisionnement) en eau peut comprendre un ensemble de capteurs pour capturer en amont lesdites données hydrométéorologiques tels que, par exemple, et de manière non limitative, des pluviomètres, des oreillers de neige ou des coussins à

30 grêlons, qui mesurent tous la quantité d'hydrométéores une fois qu'ils sont tombés sur le sol. Les oreillers de neige mesurent le poids de la neige tombée

et le convertissent en un équivalent en liquide. Les coussins à grêlons montrent la taille approximative et la quantité représentative des grêlons qui sont tombés à un emplacement. Certains capteurs peuvent mesurer la quantité d'eau distribuée quotidiennement. Un paramètre pertinent dans le processus d'alimentation en eau, par exemple, peut être la quantité d'eau disponible quotidiennement en fonction d'un ensemble d'événements hydrométéorologiques tels que décrits ci-dessus. L'infrastructure mettant en œuvre le processus industriel d'alimentation en eau peut comprendre un ensemble de capteurs capturant la quantité de l'eau disponible quotidiennement. Les données relatives audit paramètre pertinent peuvent être mémorisées dans une base de données pour le processus de prédiction conformément au procédé décrit ci-dessus. La deuxième valeur de sortie prédite de la quantité d'eau disponible quotidiennement peut être comparée à une valeur mesurée réelle de ladite quantité d'eau disponible quotidiennement. Ledit résultat de la comparaison est ainsi utilisé pour entreprendre certaines actions. Par exemple, et de manière non limitative, si la deuxième valeur de sortie prévue est supérieure à celle mesurée, une action peut ainsi être, par exemple, l'augmentation de la taille du réservoir stockant l'eau fournie. Si, au contraire, ladite deuxième valeur de sortie prévue est inférieure à celle mesurée, une action entreprise peut consister à construire un réservoir ou des cuves pour retenir, par exemple, davantage d'eau pendant les épisodes pluvieux.

Dans un autre mode de réalisation, le paramètre pertinent peut être la quantité d'eau tombée. Dans ce cas, les multiples deuxièmes valeurs de sortie prédites peuvent éviter l'inondation d'une région en les intégrant dans un système d'alerte pour entreprendre des actions concrètes, telles que la construction de digues, par exemple.

Dans certains modes de réalisation, le processus industriel peut concerner la production d'énergie dans un système hydroélectrique. Dans ce cas, les mêmes données d'historique hydrométéorologiques mémorisées peuvent être utilisées. Un paramètre pertinent est l'énergie de sortie dudit système hydroélectrique. La base de données peut également contenir des

données relatives à l'écoulement d'eau utilisé pour produire les énergies de sortie, ou des données relatives au barrage hydroélectrique. Selon que les multiples deuxièmes valeurs de sortie prédites de l'énergie hydroélectrique sont inférieures ou supérieures aux valeurs de sortie mesurées réelles de l'énergie hydroélectrique, le flux d'eau pour produire ladite énergie de sortie peut être augmenté ou réduit.

Dans certains modes de réalisation, le premier modèle de prédiction utilisé pour générer des premières valeurs de sortie prédites d'un paramètre pertinent est un modèle ARFIMA (modèle autorégressif à moyennes mobiles fractionnairement intégré).

Le modèle ARFIMA tient compte d'un continuum de différences fractionnaires $-0,5 < d < 0,5$, tandis que la généralisation du modèle autorégressif à moyennes mobiles (ARMA) classique aux différences fractionnaires permet de rechercher des ruptures structurelles et cela a été utile dans des domaines aussi variés que l'hydrologie ou l'énergie par exemple. La forme générale d'un modèle ARFIMA(p,d,q) pour modéliser la valeur de sortie quotidienne du paramètre pertinent y_t du processus industriel est donnée par

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t$$

où le paramètre d est une valeur non entière comprise entre $[-0,5;0,5]$, y_t est la valeur de sortie quotidienne du paramètre pertinent à l'instant t , ε_t est une composante stochastique normalement distribuée avec une moyenne 0 et une variance σ^2 , et $\phi(B)$ et $\theta(B)$ représentent respectivement les composantes AR (autorégressives) et MA (de moyenne mobile) avec un opérateur de retard B . Dans un modèle fractionnaire, la puissance peut être fractionnaire, avec la signification du terme identifié utilisant le développement en série binomiale formelle suivant :

$$(1-B)^d = 1 - dB + \frac{d(d-1)}{2!}B^2 - \dots$$

De plus, B^k est l'opérateur de décalage ou de retard, c'est-à-dire, qu'étant donné y_t , nous avons $B^k y_t = y_{t-k}$. La formulation de la valeur de sortie du paramètre pertinent y_t par le modèle ARFIMA indiquée par y_t^{fi} peut
 5 être réécrite en tant que :

$$y_t^{fi} = \phi^{-1}(B)(1 - B)^{-d}\theta(B)\varepsilon_t$$

Le modèle ARFIMA, ayant le paramètre d pour la technique d'ordre
 10 fractionnaire, pourrait capturer à la fois la dépendance à long terme et la dépendance à court terme. En général, les estimateurs de d peuvent être obtenus en utilisant la probabilité maximale.

Pour la prédiction des valeurs de sortie du paramètre pertinent du processus industriel, une prédiction de séries chronologiques à multiples pas
 15 (pas temporel) (également appelée à long terme) est utilisée. Ladite prédiction à multiples pas consiste, en général, à prédire les H valeurs suivantes $[y_{m+1}, \dots, y_{m+H}]$ compte tenu d'une série chronologique à une variable $[y_1, \dots, y_m]$ comprenant m observations ou données d'historique mémorisées du processus industriel ou du paramètre pertinent dudit processus industriel.

20 Dans l'approche à multiples pas (multi-step ahead), une stratégie de prédiction récursive et une stratégie de prédiction directe peuvent être utilisées. Avec la stratégie récursive, des prédictions sont générées en utilisant un modèle avec à un pas, appliqué de manière itérative pour le nombre de pas souhaité. Avec la stratégie directe, un modèle spécifique à horizon est estimé
 25 et des prédictions sont calculées directement par le modèle estimé pour chaque horizon de prédiction.

La stratégie récursive itère, H fois, un modèle de prédiction à un pas pour obtenir les H prédictions. Après l'estimation de la série future, elle est renvoyée en tant qu'entrée pour la prédiction suivante. Dans cette stratégie,
 30 un modèle unique f est entraîné pour effectuer une prédiction à un pas. Ce modèle suppose une dépendance autorégressive de la valeur future de la

série chronologique vis-à-vis des m valeurs (ordre de retard ou d'intégration) passées et d'un terme de bruit à moyenne nulle supplémentaire ω comme suit

$$y_{t+1} = f(y_t, \dots, y_{t-m+1}) + \omega$$

5

En fonction du bruit présent dans la série chronologique et de l'horizon de prédiction, la stratégie récursive peut présenter une faible performance dans les tâches de prédiction à multiples pas.

La stratégie directe (DirStr) estime un ensemble de H modèles de prédiction, renvoyant chacun une prédiction pour l'horizon $h \in [1, \dots, H]$. En d'autres termes, H modèles f_h sont appris (un pour chaque horizon) à partir des données de série chronologique mémorisées $[y_1, \dots, y_m]$ où

15

$$y_{t+h} = f_h(y_t, \dots, y_{t-m+1}) + \omega$$

avec $t \in [1, \dots, n - H]$ et $h \in [1, \dots, H]$. Les prédictions estimées $\hat{y}_{t+h}$ sont obtenues en utilisant les H modèles $\hat{f}_h$ appris comme suit

20

$$\hat{y}_{t+h} = \hat{f}_h(y_t, \dots, y_{t-m+1})$$

L'horizon de prédiction h peut être vu comme la durée dans le futur pendant laquelle la prédiction doit être faite. En résumé, la stratégie de prédiction directe est basée sur H apprenants (modèles d'apprentissage). Cela signifie un apprenant pour prédire la valeur à l'instant $t+1$, un autre à l'instant $t+2$, etc. Ainsi, par exemple, lorsque la prédiction concerne une prédiction sur cinq jours, cinq apprenants peuvent ainsi être utilisés avec la stratégie directe.

Dans certains modes de réalisation, chaque modèle de la famille de H modèles d'apprentissage automatique de prédiction $\hat{f}_h$ utilisés pour prévoir les multiples deuxièmes valeurs de sortie prédites du paramètre pertinent y_t est construit au moyen d'une approche de stratégie directe élargie (DirStrX) dans laquelle les premières valeurs de sortie prédites mémorisées y_t^{fi} sont

30

incorporées en tant que variables d'entrée exogènes en plus des données de longues séries chronologiques d'historique mémorisées relatives au paramètre pertinent du processus industriel.

$$\hat{y}_{t+h} = \hat{f}_h^X(y_t, \dots, y_{t-m+1}, y_t^{fi}, \dots, y_{t-m+1}^{fi})$$

La stratégie directe réduit à un minimum les erreurs prédites avec h pas avant au lieu des erreurs avec un pas comme pour la stratégie récursive.

La figure 3 montre, par exemple et de manière non limitative, la différence entre la stratégie directe (DirStr) et la stratégie directe étendue (DirStrX) pour un ordre de modèle $m = 3$ et un horizon de prédiction $h = 3$.

Dans certains modes de réalisation, le modèle d'apprentissage automatique $\hat{f}_h$ est un modèle ANN (réseau neuronal artificiel). Dans un autre mode de réalisation, le modèle d'apprentissage automatique $\hat{f}_h$ est un modèle k-NN (k voisins les plus proches). Dans un autre mode de réalisation, le modèle d'apprentissage automatique $\hat{f}_h$ est un modèle SVR (régression basée sur une machine à vecteur de support ou séparateurs à vaste marge).

Dans certains modes de réalisation, un procédé ANN, un procédé k-NN et un procédé SVR peuvent être construits simultanément pour la prédiction de la deuxième valeur de sortie du paramètre pertinent du processus industriel.

Dans certains modes de réalisation, le nombre minimum de jours, D, des données d'historique mémorisées pour améliorer le rendement du processus industriel est compris dans un intervalle défini par $100 < D < 300$. De préférence, le nombre minimum de jours, D, est fixé à 250 jours. Un tel nombre minimum de jours est suffisant pour capturer la mémoire longue et la persistance du phénomène de variabilité d'un paramètre pertinent du processus industriel.

L'invention concerne également un système (1) pour améliorer le rendement d'un processus industriel.

Dans certains modes de réalisation, le système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle, comme illustré sur la figure 2, comprend au moins un ensemble
5 de capteurs (10) pour capturer au moins des valeurs de sortie d'au moins un paramètre pertinent du processus industriel, une base de données (11) pour mémoriser lesdites données de sortie capturées et au moins des données de longues séries chronologiques d'historique relatives au paramètre pertinent ou au processus industriel, au moins une infrastructure informatique comprenant
10 au moins un processeur sur lequel sont exécutés un ensemble de programmes mémorisés dans une mémoire pour prédire et mémoriser au moins de multiples valeurs de sortie d'un paramètre pertinent et déclencher, automatiquement ou manuellement, un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle ledit processus
15 industriel s'exécute, au moyen de la mise en œuvre du procédé pour améliorer le rendement d'un processus industriel tel que décrit dans les sections précédentes.

Dans certains modes de réalisation, le système (1) comprend un
20 dispositif de prédiction (12) contenant au moins :

- une première unité de prédiction (120) comprenant au moins une première mémoire pour mémoriser des données, un processeur exécutant le premier modèle de prédiction pour calculer les premières valeurs de sortie prédites du paramètre pertinent et
25 les mémoriser dans la première mémoire ;
- une unité de modélisation (121) comprenant au moins une deuxième mémoire pour mémoriser des données, un processeur pour exécuter un ensemble de programmes pour construire le deuxième modèle de prédiction de H modèles d'apprentissage
30 automatique et les mémoriser dans la deuxième mémoire ;

- une deuxième unité de prédiction (122) comprenant une troisième mémoire pour mémoriser des données, un processeur pour exécuter le deuxième modèle de prédiction pour calculer les multiples deuxième valeurs de sortie prédites du paramètre pertinent et les erreurs prédites associées, et mémoriser lesdites valeurs de sortie et les erreurs prédites dans la troisième mémoire.

Dans certains modes de réalisation, le système (1) comprend un dispositif d'analyse de données (13) comprenant au moins un processeur, au moins une mémoire pour mémoriser des données et/ou un ensemble de programmes dont l'exécution sur le processeur met en œuvre la comparaison des multiples deuxième valeurs de sortie prédites du paramètre pertinent et des valeurs de sortie réelles dudit paramètre pertinent et génère au moins un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

Dans certains modes de réalisation, le dispositif d'analyse de données (13) transmet le résultat de l'analyse à une interface interactive d'un opérateur du processus industriel, ladite interface interactive comprenant au moins un agencement pour permettre à l'opérateur de générer, après avoir vérifié le résultat de l'analyse, un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

Dans certains modes de réalisation, le système (1) comprend un dispositif de déclenchement (14), connecté au moins à une unité de commande du processus industriel, contenant au moins un processeur, au moins une mémoire pour mémoriser des données et/ou des instructions exécutables sur le processeur de manière à déclencher l'ensemble d'actions mémorisé à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

Dans certains modes de réalisation, le dispositif de déclenchement (14) peut comprendre un agencement pour appliquer automatiquement l'ensemble

d'actions au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

Par exemple, le dispositif de déclenchement (14) peut être connecté à une unité de commande du processus industriel et, par exemple à un bouton d'alimentation qui peut arrêter le processus, l'agencement peut être un dispositif mécanique connecté au bouton d'alimentation. Si, après l'analyse du résultat de la prédiction de la deuxième valeur de sortie du paramètre pertinent, l'action choisie consiste à arrêter le processus industriel, le dispositif de déclenchement (14) peut actionner le dispositif mécanique qui, à son tour, enfoncera le bouton d'alimentation pour le désactiver.

Dans certains modes de réalisation, le dispositif de déclenchement (14) peut comprendre un agencement pour appliquer manuellement l'ensemble d'actions au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

Par exemple, et de manière non limitative, le dispositif de déclenchement (14) peut comprendre un ensemble de boutons qui peuvent commander le fonctionnement du processus industriel et/ou de l'infrastructure sur laquelle le processus industriel s'exécute. Ledit dispositif de déclenchement peut également comprendre une alarme pour alerter un opérateur du fait qu'un ensemble d'actions doivent être appliquées audit processus. L'opérateur peut ainsi, manuellement, pousser au moins un bouton parmi l'ensemble de boutons de manière à appliquer au moins une action parmi l'ensemble d'actions nécessaire pour surveiller le processus et/ou l'infrastructure sur laquelle le processus industriel s'exécute.

Dans certains modes de réalisation, la première unité de prédiction (120) est connectée, par l'intermédiaire de moyens de communication, à l'unité de modélisation (121) de manière à transmettre au moins les premières valeurs de sortie prédites du paramètre pertinent du processus industriel, lesdites valeurs de sortie prédites étant mémorisées dans la deuxième mémoire.

Dans certains modes de réalisation, l'unité de modélisation (121) est connectée, par l'intermédiaire de moyens de communication, à la deuxième

unité de prédiction (122) de manière à transmettre le deuxième modèle de prédiction de H modèles d'apprentissage automatique, ledit deuxième modèle de prédiction étant mémorisé dans la troisième mémoire.

Dans certains modes de réalisation, la première unité de prédiction (120) est connectée, par l'intermédiaire de moyens de connexion, à la deuxième unité de prédiction (122) de manière à transmettre les premières valeurs de sortie prédites du paramètre pertinent du processus industriel utilisées en tant que variables d'entrée pour le deuxième modèle de prédiction, lesdites valeurs de sortie prédites étant mémorisées dans la troisième mémoire.

Dans certains modes de réalisation, la deuxième unité de prédiction (122) est connectée, par l'intermédiaire de moyens de communication, au dispositif d'analyse de données (13) de manière à transmettre au moins les multiples deuxième valeurs de sortie prédites du paramètre pertinent et les erreurs prédites associées auxdites multiples deuxième valeurs de sortie prédites.

Dans certains modes de réalisation, le dispositif d'analyse de données (13) est connecté, par l'intermédiaire de moyens de communication, au dispositif de déclenchement (14) de manière à transmettre un résultat de l'analyse, et l'ensemble d'actions à appliquer.

Par exemple, et de manière non limitative, en ce qui concerne les systèmes hydroélectriques, le dispositif de déclenchement (14) peut être connecté au système surveillant le flux d'eau dans les turbines, qui convertissent l'énergie du flux d'eau en une rotation mécanique de manière à actionner des alternateurs pour produire l'énergie de sortie. Lors de la réception des instructions ou des informations contenant au moins l'ensemble d'actions à appliquer du dispositif d'analyse de données (13), ledit dispositif de déclenchement (14) peut actionner automatiquement l'ouverture du système contrôlant l'entrée d'eau dans les turbines de manière à réduire ou augmenter le flux d'eau dans les turbines.

Dans certains modes de réalisation, la première unité de prédiction (120), la deuxième unité de prédiction (122), l'unité de modélisation (121) et

au moins le dispositif d'analyse de données (13) sont connectés à la base de données (11) de manière à récupérer des données d'historique et/ou des données réelles relatives au processus industriel ou au paramètre pertinent dudit processus industriel.

5 Dans certains modes de réalisation, les moyens de communication de la première unité de prédiction (120), de la deuxième unité de prédiction (122), de l'unité de modélisation (121), du dispositif d'analyse de données (13) et au moins du dispositif de déclenchement (14) consistent en des moyens de communication sans fil/filaires (Wifi, USB, Ethernet, etc.).

10

L'invention concerne également un produit programme d'ordinateur pour améliorer le rendement d'un processus industriel.

15 Dans certains modes de réalisation, le produit programme d'ordinateur est mémorisé sur un support pouvant être lu par des moyens informatiques et contenant des instructions exécutables par au moins un processeur de ces moyens informatiques pour mettre en œuvre le procédé tel que décrit ci-dessus.

20 Dans ce qui suit, nous présentons un exemple pour la compréhension du principe de la présente invention et particulièrement de son efficacité. Dans l'exemple, le paramètre pertinent peut ne pas être un paramètre physique, cependant, on doit comprendre que le but dudit exemple est simplement d'illustrer comment le système de la présente invention fonctionne.

25 Par exemple, et de manière non limitative, le paramètre à prédire dans cet exemple est le rendement composé en continu ou le rendement logarithmique (également appelé simplement rendement). Le but est d'évaluer la performance du procédé présenté dans la présente demande.

30 Dans la modélisation de série chronologique financière, par exemple, les quantités suivantes présentent un intérêt et sont souvent utilisées pour définir le rendement ou la volatilité : $P_t^{(o)}$, $P_t^{(c)}$, $P_t^{(h)}$, $P_t^{(l)}$ qui sont,

respectivement, les cours des actions à l'ouverture, à la fermeture du jour de trading et les valeurs maximum et minimum pour chaque jour de trading. La disponibilité d'un ensemble d'apprentissage de T cours historiques est supposée. Le rendement pour le jour t calculé à partir des cours à la fermeture

5 $P_t^{(c)}$ et $P_{t-1}^{(c)}$ est donné par

$$r_t = \ln(P_t^{(c)} / P_{t-1}^{(c)})$$

Un premier estimateur σ_t^0 du proxy de volatilité peut être défini comme
 10 la valeur au carré du rendement tel que défini ci-dessus. Afin d'optimiser l'estimation dudit proxy de volatilité, une famille de proxies σ_t^i peut être déduite en incorporant des informations supplémentaires telles que les cours à l'ouverture, maximum et minimum pour un jour de trading donné en plus du cours à la fermeture. Dans ce cas, une estimation de la volatilité peut être
 15 donnée par

$$\sigma_t^1 = 0.511(u - d)^2 - (2 \ln 2 - 1)c^2$$

avec $u = \ln(P_t^{(h)} / P_{t-1}^{(o)})$, $d = \ln(P_t^{(l)} / P_{t-1}^{(o)})$, $c = \ln(P_t^{(c)} / P_{t-1}^{(o)})$ et où u, d et
 20 c sont respectivement le cours élevé normalisé, le cours bas normalisé et le cours à la fermeture normalisé.

Afin d'évaluer la performance de prédiction du procédé décrit dans la présente demande, une comparaison sera faite avec un modèle ARMA-GARCH(P,Q) hybride (appelé ici GarchX), un modèle kNN, un modèle SVR,
 25 un modèle ANN et les modèles combinés ARFIMA-kNN, ARFIMA-NN et ARFIMA-SVR.

Dans le modèle GarchX, le rendement est généralement donné par

$$r_t^{ag} = \mu + \sum_{i=1}^P a_i r_{t-i} + \sum_{j=1}^Q b_j \varepsilon_{t-j} + \varepsilon_t$$

- avec une moyenne μ , des coefficients autorégressifs a_i et des coefficients de moyenne mobile b_j . Le terme ε_t est une composante stochastique avec la forme $\varepsilon_t = \sigma_t z_t$ ($z_t \sim N(0,1)$) où les z_t sont des variables aléatoires indépendantes et identiquement distribuées avec une moyenne et
- 5 une variance unitaire, et σ_t est formulé en utilisant un prédicteur externe, σ_t^1 représentant ledit prédicteur externe

$$\sigma_t = \sigma_t^{Gx} = \sqrt{(\omega + \zeta \sigma_t^1) + \sum_{j=1}^P \beta_j (\sigma_{t-j})^2 + \sum_{i=1}^Q \alpha_i \varepsilon_{t-i}^2}$$

- 10 Les coefficients $\omega, \alpha_i, \beta_j, \zeta$ sont ajustés conformément à une procédure à probabilité logarithmique maximum (ou la fonction logarithmique de la probabilité) estimée. Dans cette étude, $P=Q=1$. Généralement, une distribution d'erreurs à asymétrie généralisée est utilisée lors de l'estimation et de la prédiction de modèles GARCH.
- 15 Pour de nombreux traders et analystes, la direction du marché est plus importante que la valeur de la prédiction elle-même, étant donné que l'argent des marchés financiers peut être réalisé simplement en connaissant la direction dans laquelle les rendements iront. Le modèle de trading utilise les valeurs précédentes des rendements en tant que caractéristiques d'entrée et
- 20 la direction du marché des jours suivants en tant que caractéristique cible. Trois stratégies de trading différentes sont définies : une stratégie « naïve », une stratégie « acheter et conserver » et une stratégie de trading « hybride ». Pour comparer les différentes stratégies de trading, plusieurs mesures de performances de trading sont utilisées telles que le ratio de Sharpe, le
- 25 rendement et la volatilité annualisés et le rapport gain/perte moyen. Ces mesures de performances de trading sont analysées en examinant les périodes pré- et post-crise financière.

- La stratégie de trading « naïve » est simpliste. En effet, une action est achetée au marché d'ouverture suivant, si le jour actuel t est haussier, c'est-
- 30 à-dire, à la hausse où le cours de fermeture $P_t^{(c)}$ du jour actuel est supérieur à

son cours d'ouverture $P_t^{(o)}$. Dans le trading, les investisseurs achètent une action ou continuent s'ils pensent que sa valeur va augmenter. De cette manière, ils peuvent la vendre à une valeur plus élevée que celle payée et récoltent un profit.

- 5 Par la même occasion, l'action est vendue à l'ouverture suivante, si le jour actuel t est baissier, c'est-à-dire, à la baisse où le cours de fermeture $P_t^{(c)}$ du jour actuel est inférieur au cours d'ouverture $P_t^{(o)}$. Dans ce cas, les investisseurs vendent d'abord dans l'espoir que la valeur du cours de l'action diminuera et qu'ils seront capables de racheter le bien ultérieurement à un
- 10 cours plus bas.

- La stratégie « acheter et conserver » est une stratégie d'investissement où un investisseur achète des actions et les conserve longtemps, avec pour but que la valeur des actions augmentera graduellement sur une longue période de temps. Celle-ci est basée sur l'idée qu'à long terme les marchés
- 15 financiers donnent un bon taux de rendement même en prenant en compte un degré de volatilité.

- La stratégie « acheter et conserver » dit que les investisseurs ne verront jamais ces rendements s'ils liquident après une baisse. Ce point de vue considère que la détermination du moment propice, c'est-à-dire, le concept
- 20 qu'on peut entrer dans le marché à prix bas et vendre à prix hauts, ne fonctionne pas ; tenter cette détermination du moment propice donne des résultats négatifs, au moins pour les petits investisseurs et les investisseurs sans grande expérience, aussi il vaut mieux qu'ils achètent et conservent simplement.

- 25 Dans la stratégie de trading « hybride », la seule différence comparée à la stratégie « naïve », est qu'on considère la tendance à terme du rendement prédit. La direction du marché du jour suivant et la position à prendre sont déterminées comme suit : si le rendement prédit $\hat{r}_t$ pour le jour suivant t est positif, acheter au marché d'ouverture le jour suivant t ; autrement, s'il est
- 30 négatif, vendre au marché d'ouverture le jour suivant t . Toutes les positions sont liquidées au marché de fermeture.

Les rendements prédits sont les revenus basés sur les modèles d'apprentissage utilisant soit simplement GarchX, KNN et SVR, soit ARFIMA-KNN, ARFIMA-SVR et ARFIMA-NN combinés. Les rendements prédits sont générés par deux catégories de stratégies de prédiction : réursive et directe

5 décrites ci-dessus, avec une approche de validation croisée de série chronologique. Cette approche implique l'apprentissage du modèle sur une fenêtre de données et la prédiction du revenu de la période suivante, et ensuite le décalage de la fenêtre d'apprentissage en avant d'une période. Le modèle est ensuite réappris sur la nouvelle fenêtre et le revenu de la période suivante

10 est prédit. Ce processus est répété sur la longueur de la série chronologique. La performance de la validation croisée du modèle est simplement la performance des prédictions du jour suivant utilisant certaines mesures de performance bien connues présentées ci-dessous.

Pour comparer la précision et la performance de trading, tous les

15 modèles sont maintenus avec une période hors échantillonnage identique. Cette étude utilise 4 métriques statistiques généralement utilisées telles que l'erreur quadratique moyenne (MSE), l'erreur absolue moyenne (MAE) et la racine de l'erreur quadratique moyenne (RMSE). Plus les valeurs de ces mesures sont petites, plus les valeurs de rendement prédites sont proches des

20 valeurs réelles.

On observe que ces mesures de précision statistiques ne sont pas suffisantes pour analyser les résultats concernant le critère financier. L'erreur de prédiction peut avoir été réduite à un minimum pendant l'estimation du modèle, mais l'évaluation du mérite réel devrait être basée sur la performance

25 des stratégies de trading. Certaines des mesures de performance plus importantes utilisées comprennent le rendement annualisé, le ratio de Sharpe, et le rapport gain/perte moyen. Le ratio de Sharpe est une mesure ajustée au risque du rendement, les ratios plus élevés étant préférés aux ratios plus faibles. Le rapport gain/perte moyen est une mesure du gain global, pour

30 lequel une valeur au-dessus de 1 est préférée. Nous examinons le rendement prédit $\hat{r}_t$ à l'instant t , n est le nombre de prédictions et $\bar{r}_n$ représente la moyenne des rendements sur la période $\{1, \dots, n\}$.

Le rendement annualisé est donné par

$$R^A = 255R^c$$

5 où R^c , le rendement cumulatif, est donné par

$$R^c = \frac{1}{n} \sum_{t=1}^n \hat{r}_t$$

Le ratio de Sharpe est donné par

10

$$Sharpe\ ratio = \frac{R^A}{\sigma^A}$$

où σ^A , la volatilité annualisée, est donné par

15

$$\sigma^A = \sqrt{255} \frac{1}{n-1} \sum_{t=1}^n (\hat{r}_t - \bar{\hat{r}}_n)^2$$

Le rapport de profit est donné par

$$Profit\ ratio = \frac{\sum_{t=1}^n \hat{r}_t > 0}{|\sum_{t=1}^n \hat{r}_t < 0|}$$

20

Le rapport de profit est le rapport entre la somme des rendements $\hat{r}_t$ positifs et la somme des rendements $\hat{r}_t$ négatifs sur la période prédite $t \in \{1, \dots, n\}$.

25 L'étude est basée sur les cours de fermeture quotidiens $P_t^{(c)}$ pour les indices du CAC40 et du S&P500. Les deux séries couvrent la période du 1^{er} janvier 2000 au 16 octobre 2017 totalisant 17 années avec 4476 jours de négociation. Ces séries ont été transformées en rendements r_t comme l'équation définie ci-dessus du rendement composé. Pour évaluer

notre étude du calme à l'agitation, la série a été divisée en deux périodes : une historique de 9 années du 1^{er} janvier 2000 au 31 décembre 2009 totalisant 2515 jours de négociation et couvrant une volatilité très faible avec une grande difficulté pendant la crise ; et une historique de 5 années du 1^{er} janvier 2012 au 16 octobre 2017, la période post-crise totalisant 1457 jours de négociation.

Pour comparer la performance des modèles de prédiction proposés, il est nécessaire de les évaluer sur des données non vues précédemment. Cette situation est probablement la plus proche d'une situation de prédiction ou de négociation vraie. Pour réaliser cela, tous les modèles ont été maintenus avec une période hors échantillonnage identique permettant une comparaison directe de leur précision de prédiction et de leur performance de trading.

Pour chaque modèle de prédiction, nous avons effectué des prédictions de rendement à pas d'un jour ou directes à pas de deux jours sur la base d'une stratégie à origine roulante avec une fenêtre mobile fixée à 250 jours de négociation en tant qu'échantillon de test. La performance moyenne est examinée parmi 30 exécutions indépendantes en débutant avec une fenêtre de 250 jours de trading prise de manière aléatoire dans les deux premières années des échantillons de données. Les différents modèles de prédiction GarchX, KNN, SVR et les modèles combinés ARFIMA-KNN, ARFIMA-SVR et ARFIMA-NN sont examinés pour tester les impacts sur la performance de trading et leur évaluation comparés à la référence de la stratégie « acheter et conserver ».

Les mesures d'erreurs utilisées sont l'erreur absolue moyenne, l'erreur quadratique moyenne, et la racine de l'erreur quadratique moyenne. Les mesures de performance de trading sont le rendement annualisé, la volatilité annualisée, le ratio de Sharpe et le rapport de profit tels que définis ci-dessus.

Nous commençons par analyser et examiner le résultat pendant la période couvrant la crise.

La figure 4a montre les vrais rendements de l'indice CAC40 sur la période couvrant la crise de 2002 jusqu'à 2010. De la même manière, la figure 5a montre les vrais rendements de l'indice S&P500 sur la période couvrant la crise de 2002 jusqu'à 2010. Sur ces figures, les procédés NN, ARFIMA-NN,

GarchX et ARFIMA-KNN montrent davantage de robustesse pendant la perturbation extrême pendant 2008-2009, tandis que les stratégies « acheter et conserver » et KNN présentent des pertes. Les tableaux 1 et 2, respectivement des figures 4b et 5b, donnent les détails des mesures d'erreurs et des performances de trading pour tous les procédés. Les trois meilleurs modèles sont NN-DirStr, ARFIMA-NN-DirStrX, GarchX. Dans le cas du CAC40, ARFIMA-NN-DirStrX et NN-DirStr montrent une haute performance avec un rendement annualisé d'environ 46 %, et un ratio de Sharpe supérieur à 1,90 et un rapport de profit supérieur à 44 %, tandis que la volatilité annualisée est similaire autour de 23 %. Pour le S&P500, avec une volatilité annualisée autour de -0,3 % comparé au CAC40, NN-DirStr, ARFIMA-NN-DirStrX donnent tous deux un bon résultat de performance. On peut observer nettement, pour les deux indices, que le modèle hybride ARFIMA-KNN-DirStrX est plus performant que le simple KNN-DirStrX étant donné que, pour l'indice CAC40, le rendement prédit annualisé passe de -1,8 % à 22,36 %, le ratio de Sharpe passe de -0,08 à 0,96, et le rapport de profit passe de -0,015 % à 19,53 %. Et pour l'indice S&P500, le rendement annualisé passe de -15 % à 7 %, le ratio de Sharpe passe de -0,67 à 0,32, et le rapport de profit passe de -0,12 % à 0,06 %.

Enfin, tous les modèles proposés, à l'exception du KNN-DirStr, capturent bien l'effet asymétrique de la volatilité et la haute perturbation pendant la crise. Lorsqu'ils sont appliqués à l'indice CAC40, le rendement prédit annualisé va de 9,9 % à 46,25 %, avec un ratio de Sharpe de 0,42 à 1,99 et un rapport de profit de 0,08 % à 45,11 %, tandis que la stratégie « acheter et conserver » donne un rendement prédit annualisé de 2,5 % avec un ratio de Sharpe de 0,09 et un rapport de profit de 0,02. Lorsqu'ils sont appliqués à l'indice S&P500, le rendement prédit annualisé va de 1,5 % à 32 %, avec un ratio de Sharpe de 0,07 à 1,41 et un rapport de profit de 0,01 % à 331,2 %, tandis que la stratégie « acheter et conserver » donne un rendement annualisé de -0,3 % avec un ratio de Sharpe de -0,01 et un rapport de profit de -0,003 %.

Lors de l'analyse et de l'examen du résultat pendant la période post-crise, on peut voir que, similairement, la figure 6a montre les vrais rendements de l'indice CAC40 sur la période post-crise de 2015 jusqu'à 2017. De la même manière, la figure 7a montre les vrais rendements de l'indice S&P500 sur la période post-crise de 2015 jusqu'à 2017. Sur ces figures, les procédés NN, ARFIMA-NN, GarchX et ARFIMA-KNN montrent de plus hautes performances comparés à la stratégie « acheter et conserver » et à KNN. Sur les tableaux 3 et 4, respectivement des figures 6b et 7b, les trois meilleurs modèles sont NN-DirStr, ARFIMA-NN-DirStrX et ARFIMA-KNN-DirStrX lorsqu'ils sont appliqués à l'indice CAC40, alors que la volatilité annualisée est d'environ 19,5 %. Dans le cas de l'indice S&P500, le tableau 4 montre que seuls ARFIMA-NN-DirStrX et NN-DirStr sont plus performants que la stratégie « acheter et conserver », alors que la volatilité annualisée est similaire autour de 13 %. Cela s'explique par le fait que la volatilité annualisée de l'indice S&P500 est plus faible comparé à l'indice CAC40, illustrant une période calme où une simple stratégie « acheter et conserver » est couronnée de succès. Dans le cas de l'indice CAC40, davantage de perturbation est remarquée et capturée par les trois meilleurs modèles identifiés, et en première place le procédé ARFIMA-NN-DirStr.

La même observation est faite pour l'indice CAC40 avec une volatilité et une perturbation plus élevées, le modèle hybride ARFIMA-KNN-DirStrX est nettement plus performant que le simple modèle KNN-DirStrX étant donné que, pour l'indice CAC40, le rendement prédit annualisé passe de 1,6 % à 20,2 %, le ratio de Sharpe passe de 0,08 à 1,04, et le rapport de profit passe de 0,02 % à 20,3 %.

Cette étude rapporte une tâche empirique qui recherche comment la performance de prédiction et de trading pourrait être améliorée en utilisant des modèles hybrides ARFIMA-apprentissage automatique couplés à une prédiction de stratégie directe à multiples pas. Ces modèles hybrides sont comparés avec de simples modèles d'apprentissage automatique utilisant une prédiction de stratégie directe, et également le modèle Garch classique avec un prédictor externe (GarchX) basé sur une stratégie récursive. La

performance des simulations de trading basées sur ces modèles est comparée avec la stratégie « acheter et conserver » bien connue. De plus, la performance de trading est estimée dans les périodes de crise lorsque le niveau de volatilité peut augmenter de façon spectaculaire en quelques jours.

5 Nous avons montré que, pendant les périodes de forte perturbation, la plupart des modèles examinés battent la stratégie « acheter et conserver » et réussissent à capturer le changement structurel avec l'effet de mémoire longue. A partir de l'analyse des résultats expérimentaux, il est nettement évident que les modèles proposés ARFIMA-NN-DirStrX, NN-DirStr et GarchX
10 fournissent un meilleur rapport de profit avec une plus grande robustesse face aux agitations. Nous avons également remarqué l'avantage du modèle ARFIMA dans le modèle hybride ARFIMA-KNN-DirStrX étant donné qu'il augmente fortement le rapport de profit comparé au simple modèle KNN-DirStr.

15 L'analyse montre également qu'il est difficile de battre la stratégie « acheter et conserver » dans une période calme et stable, caractérisée par une faible volatilité telle que rencontrée avec l'indice S&P500 dans la période post-crise.

20 Il a été montré, par cet exemple, combien le procédé décrit dans la présente demande peut être rentable.

Un autre exemple d'application du présent procédé et/ou du présent système, de manière non limitative, est la prédiction de l'intensité de la vitesse du vent pour la gestion des risques météorologiques. En effet, les producteurs
25 d'électricité sont intéressés par les prédictions à long terme de la vitesse du vent pour au moins deux raisons principales :

- . pour évaluer la rentabilité de la construction d'une ferme éolienne à un emplacement donné ; et
- . pour compenser les risques associés à la variabilité de la vitesse du
30 vent pour une ferme éolienne déjà en fonctionnement.

Le procédé, tel que décrit dans la présente demande, peut être utilisé pour générer avec plus de précision une prédiction de l'évolution à long terme

de l'intensité de la vitesse du vent (qui est, dans ce cas, le paramètre pertinent pour le processus de production d'électricité) et pourrait être utilisé pour alerter à temps de la nécessité d'entreprendre des actions contre un risque de dommages dus à une intensité de la vitesse du vent inattendue.

5

La présente demande décrit diverses caractéristiques techniques et avantages en référence aux figures et/ou à divers modes de réalisation. L'homme de métier comprendra que les caractéristiques techniques d'un mode de réalisation donné peuvent en fait être combinées avec des caractéristiques d'un autre mode de réalisation à moins que l'inverse ne soit explicitement mentionné ou qu'il ne soit évident que ces caractéristiques sont incompatibles ou que la combinaison ne fournisse pas une solution à au moins un des problèmes techniques mentionnés dans la présente demande. De plus, les caractéristiques techniques décrites dans un mode de réalisation donné peuvent être isolées des autres caractéristiques de ce mode à moins que l'inverse ne soit explicitement mentionné.

Il doit être évident pour les personnes versées dans l'art que la présente invention permet des modes de réalisation sous de nombreuses autres formes spécifiques sans l'éloigner du domaine d'application de l'invention comme revendiqué. Par conséquent, les présents modes de réalisation doivent être considérés à titre d'illustration, mais peuvent être modifiés dans le domaine défini par la protection demandée, et l'invention ne doit pas être limitée aux détails donnés ci-dessus.

REVENDEICATIONS

1. Procédé (2) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle comprenant au moins une base de données comprenant des données de longues séries chronologiques d'historique du processus industriel ou d'au moins un paramètre pertinent dudit processus industriel, au moins une infrastructure informatique contenant au moins une mémoire pour mémoriser un ensemble de programmes, au moins un processeur pour exécuter ledit procédé comprenant au moins une étape consistant à :
- Générer, via un dispositif de prédiction, des premières valeurs de sortie prédites (20) d'un paramètre pertinent du processus industriel avec un premier modèle de prédiction mémorisé et les erreurs associées, sur la base des données de longues séries chronologiques d'historique relatives audit paramètre pertinent mémorisées dans la base de données,
- le procédé étant caractérisé en ce qu'il comprend également au moins les étapes suivantes consistant à :
- construire (21), via le dispositif de prédiction, une famille de H modèles d'apprentissage automatique (ML) de prédiction, constituant le deuxième modèle de prédiction, pour la prédiction de multiples valeurs de sortie du même paramètre pertinent sur la base des données de séries chronologiques d'historique mémorisées et des premières valeurs de sortie prédites mémorisées et générer, avec ledit deuxième modèle de prédiction, de multiples deuxième valeurs de sortie prédites (22) dudit paramètre pertinent et les erreurs associées ;
 - comparer (23), via un dispositif d'analyse de données, les multiples deuxième valeurs de sortie prédites du paramètre pertinent et les valeurs de sortie réelles mémorisées dudit paramètre pertinent ;
 - déterminer (24), via le dispositif d'analyses de données, l'ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur

laquelle le processus industriel s'exécute, sur la base de l'analyse de comparaison, et mémoriser lesdites actions ;

- déclencher (25) automatiquement ou manuellement, via au moins un dispositif de déclenchement, l'ensemble d'actions mémorisé et surveiller la sortie du paramètre pertinent sur la base au moins des erreurs prédites obtenues à partir de la deuxième étape de prédiction.

2. Procédé pour améliorer le rendement d'un processus industriel selon la revendication 1, caractérisé en ce que le premier modèle de prédiction utilisé pour générer des premières valeurs de sortie prédites (20) d'un paramètre pertinent est un modèle ARFIMA (modèle autorégressif à moyennes mobiles fractionnairement intégré).

3. Procédé pour améliorer le rendement d'un processus industriel selon la revendication 1, caractérisé en ce que chaque modèle de la famille de H modèles d'apprentissage automatique de prédiction utilisé pour prévoir les multiples deuxièmes valeurs de sortie prédites du paramètre pertinent est construit (21) au moyen d'une approche de stratégie directe élargie dans laquelle les premières valeurs de sortie prédites mémorisées sont incorporées en tant que variables d'entrée exogènes en plus des données de longues séries chronologiques d'historique mémorisées relatives au paramètre pertinent du processus industriel.

4. Procédé pour améliorer le rendement d'un processus industriel selon la revendication 1, caractérisé en ce que le modèle d'apprentissage automatique est un modèle ANN (réseau neuronal artificiel).

5. Procédé pour améliorer le rendement d'un processus industriel selon la revendication 1, caractérisé en ce que le modèle d'apprentissage automatique est un modèle k-NN (k voisins les plus proches).

6. Procédé pour améliorer le rendement d'un processus industriel selon la revendication 1, caractérisé en ce que le modèle d'apprentissage automatique est un modèle SVR (régression basée sur une machine à vecteur de support).

7. Procédé selon les revendications 1 à 6, caractérisé en ce que le nombre de jour minimum, D, des données d'historique mémorisées pour

améliorer le rendement du processus industriel est compris dans un intervalle défini par $100 < D < 300$.

8. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle, le système (1) comprenant au moins un ensemble de capteurs pour capturer au moins des valeurs de sortie d'au moins un paramètre pertinent du processus industriel, une base de données (11) pour mémoriser lesdites données de sortie capturées et au moins des données de longues séries chronologiques d'historique relatives au paramètre pertinent ou au processus industriel, au moins une infrastructure informatique comprenant au moins un ensemble de dispositifs comportant au moins un processeur sur lequel sont exécutés un ensemble de programmes mémorisés dans une mémoire pour prédire et mémoriser au moins de multiples valeurs de sortie d'un paramètre pertinent et déclencher, automatiquement ou manuellement, un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle ledit processus industriel s'exécute, au moyen de la mise en œuvre du procédé pour améliorer le rendement d'un processus industriel selon la revendication 1.

9. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 8, caractérisé en ce qu'il comprend un dispositif de prédiction (12) contenant au moins :

- une première unité de prédiction (120) comprenant au moins une première mémoire pour mémoriser des données, un processeur exécutant le premier modèle de prédiction pour calculer les premières valeurs de sortie prédites du paramètre pertinent et les mémoriser dans la première mémoire ;
- une unité de modélisation (121) comprenant au moins une deuxième mémoire pour mémoriser des données, un processeur pour exécuter un ensemble de programmes pour construire le deuxième modèle de prédiction de H modèles d'apprentissage automatique et les mémoriser dans la deuxième mémoire ;

- une deuxième unité de prédiction (122) comprenant une troisième mémoire pour mémoriser des données, un processeur pour exécuter le deuxième modèle de prédiction pour calculer les multiples deuxièmes valeurs de sortie prédites du paramètre pertinent et les erreurs prédites associées, et mémoriser lesdites valeurs de sortie et erreurs prédites dans la troisième mémoire.

10. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 8, caractérisé en ce qu'il comprend un dispositif d'analyse de données (13) comprenant au moins un processeur, au moins une mémoire pour mémoriser des données et/ou un ensemble de programmes dont l'exécution sur le processeur met en œuvre la comparaison des multiples deuxièmes valeurs de sortie prédites du paramètre pertinent et des valeurs de sortie réelles dudit paramètre pertinent et génère au moins un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

11. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 10, caractérisé en ce que le dispositif d'analyse de données (13) transmet le résultat de l'analyse à une interface interactive d'un opérateur du processus industriel, ladite interface interactive comprenant au moins un agencement pour permettre à l'opérateur de générer, après avoir vérifié le résultat de l'analyse, un ensemble d'actions à appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

12. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 8, caractérisé en ce qu'il comprend un dispositif de déclenchement (14), connecté à au moins une unité de commande du processus industriel, contenant au moins un processeur, au moins une mémoire pour mémoriser des données et/ou des instructions exécutables sur le processeur de manière à déclencher l'ensemble d'actions mémorisé à

appliquer au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

13. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 12, caractérisé en ce que le dispositif de déclenchement (14) comprend un agencement pour appliquer automatiquement l'ensemble d'actions au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

14. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 12, caractérisé en ce que le dispositif de déclenchement (14) comprend un agencement pour appliquer manuellement l'ensemble d'actions au processus industriel et/ou à l'infrastructure sur laquelle le processus industriel s'exécute.

15. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 12, caractérisé en ce que le dispositif de déclenchement (14) comprend également une alarme pour alerter un opérateur du fait qu'un ensemble d'actions doivent être appliquées audit processus.

16. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 9, caractérisé en ce que la première unité de prédiction (120) est connectée, par l'intermédiaire de moyens de communication, à l'unité de modélisation (121) de manière à transmettre au moins les premières valeurs de sortie prédites du paramètre pertinent du processus industriel, lesdites valeurs de sortie prédites étant mémorisées dans la deuxième mémoire.

17. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 9, caractérisé en ce que l'unité de modélisation (121) est connectée, par l'intermédiaire de moyens de communication, à la deuxième unité de prédiction (122) de manière à transmettre le deuxième modèle de

prédiction de H modèles d'apprentissage automatique, ledit deuxième modèle de prédiction étant mémorisé dans la troisième mémoire.

18. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 9, caractérisé en ce que la première unité de prédiction (120) est connectée, par l'intermédiaire de moyens de connexion, à la deuxième unité de prédiction (122) de manière à transmettre les premières valeurs de sortie prédites du paramètre pertinent du processus industriel utilisées en tant que variables d'entrée pour le deuxième modèle de prédiction, lesdites valeurs de sortie prédites étant mémorisées dans la troisième mémoire.

19. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 9, caractérisé en ce que la deuxième unité de prédiction (122) est connectée, par l'intermédiaire de moyens de communication, au dispositif d'analyse de données (13) de manière à transmettre au moins les multiples deuxième valeurs de sortie prédites du paramètre pertinent et les erreurs prédites relatives auxdites multiples deuxième valeurs de sortie prédites.

20. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 10, caractérisé en ce que le dispositif d'analyse de données (13) est connecté, par l'intermédiaire de moyens de communication, au dispositif de déclenchement (14) de manière à transmettre un résultat de l'analyse, et l'ensemble d'actions à appliquer.

21. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon la revendication 9, caractérisé en ce que la première unité de prédiction (120), la deuxième unité de prédiction (122), l'unité de modélisation (121) et au moins le dispositif d'analyse de données (13) sont connectés à la base de données de manière à récupérer les données d'historique et/ou les données réelles relatives au processus industriel ou au paramètre pertinent dudit processus industriel.

22. Système (1) pour améliorer le rendement d'un processus industriel dans une infrastructure de production industrielle selon les revendications 16 à 20, caractérisé en ce que les moyens de communication de la première unité de prédiction (120), de la deuxième unité de prédiction (122), de l'unité de modélisation (121), du dispositif d'analyse de données (13) et au moins du dispositif de déclenchement (14) sont des moyens de communication sans fils/filaires.

23. Produit programme d'ordinateur, mémorisé sur un support pouvant être lu par des moyens informatiques et contenant des instructions exécutables par au moins un processeur de ces moyens informatiques pour mettre en œuvre le procédé selon la revendication 1.

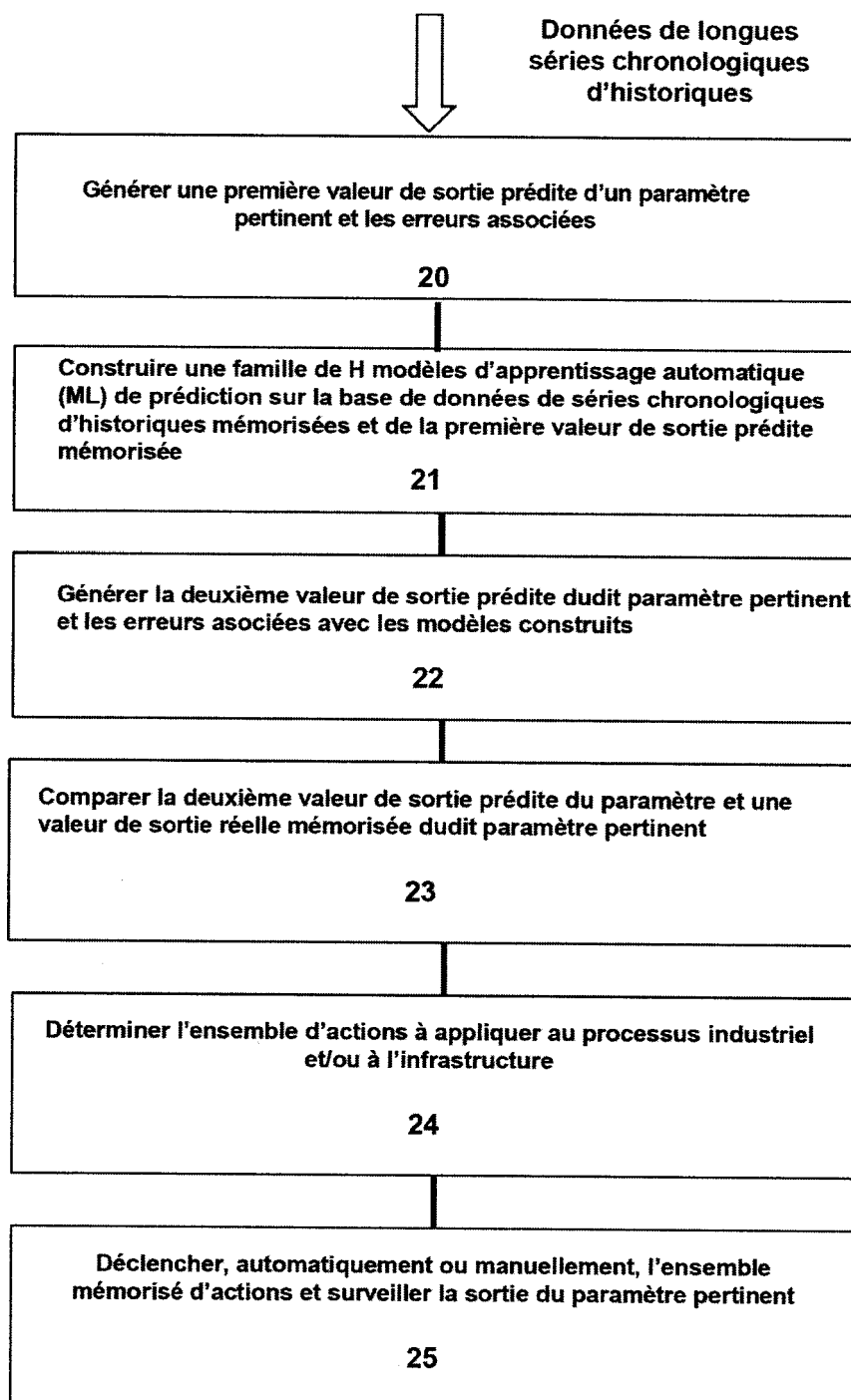


Figure 1

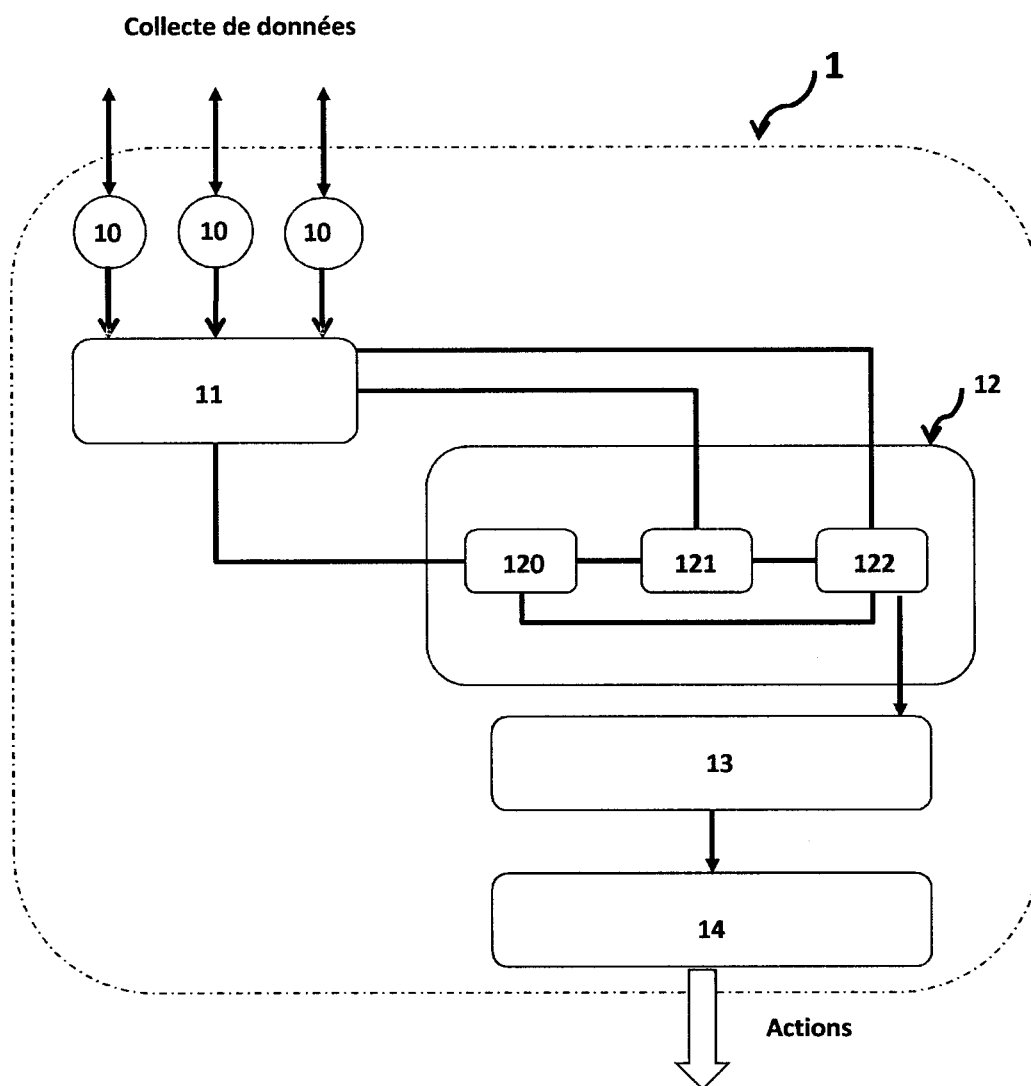


Figure 2

DirStr

- Un modèle unique f^h pour chaque horizon h .
- La prévision au pas h est effectuée en utilisant le h^e modèle.

r	$\hat{r}$
r_3	$\hat{r}_1$
r_4	$\hat{r}_2$
$\dots$	$\dots$
r_{t-5}	$\hat{r}_{t-7}$
r_{t-6}	$\hat{r}_{t-2}$

DirStrX

- Un modèle unique f_h^x pour chaque horizon h .
- La prévision au pas h est effectuée en utilisant le h^e modèle.

r					$\hat{r}$
r_3	r_1	$r_3^{f_1}$	$r_2^{f_1}$	$r_1^{f_1}$	$\hat{r}_5$
r_4	r_3	r_2	$r_4^{f_2}$	$r_3^{f_2}$	$\hat{r}_6$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
r_{t-5}	r_{t-6}	r_{t-7}	$r_{t-5}^{f_t}$	$r_{t-6}^{f_t}$	$\hat{r}_{t-2}$

Figure 3

4/11

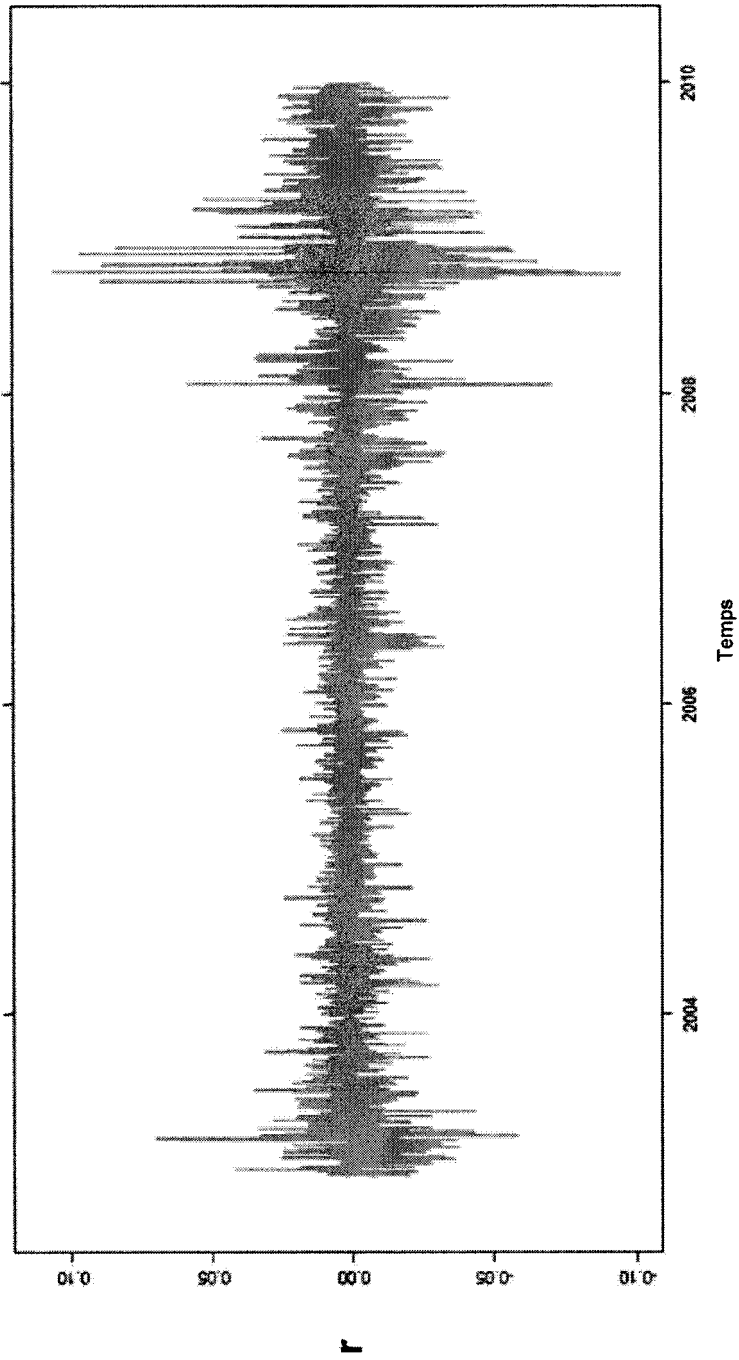


Figure 4a

Stratégies de modèles	MSE	MAE	RMSE	R^A	σ^A	Ratio de Sharpe	nUP	nDown	Rapport de profit
GarchX	0.00021	0.00992	0.01468	20.24%	23.34%	0.87	950	842	17.52%
NN-DirStr	0.00020	0.00938	0.01437	46.25%	23.19%	1.99	1085	707	45.11%
NN-DirStrX	0.00020	0.00929	0.01420	45.50%	23.20%	1.96	1078	714	44.21%
KNN-DirStr	0.00023	0.01038	0.01512	-1.8%	23.37%	-0.08	875	917	-0.015%
KNN-DirStrX	0.00022	0.01005	0.01475	22.36%	23.33%	0.96	930	862	19.53%
SVR-DirStr	0.00021	0.00997	0.01463	14.3%	23.36%	0.61	922	870	12.06%
SVR-DirStrX	0.00022	0.01001	0.01482	9.9%	23.37%	0.42	928	866	0.08%
Naïve	0.00050	0.0108	0.02256	-4%	23.37%	-0.16	872	922	-0.03%
Acheter et conserver	-	-	-	2.5%	25.18%	0.09	932	862	0.02%

Figure 4b

6/11

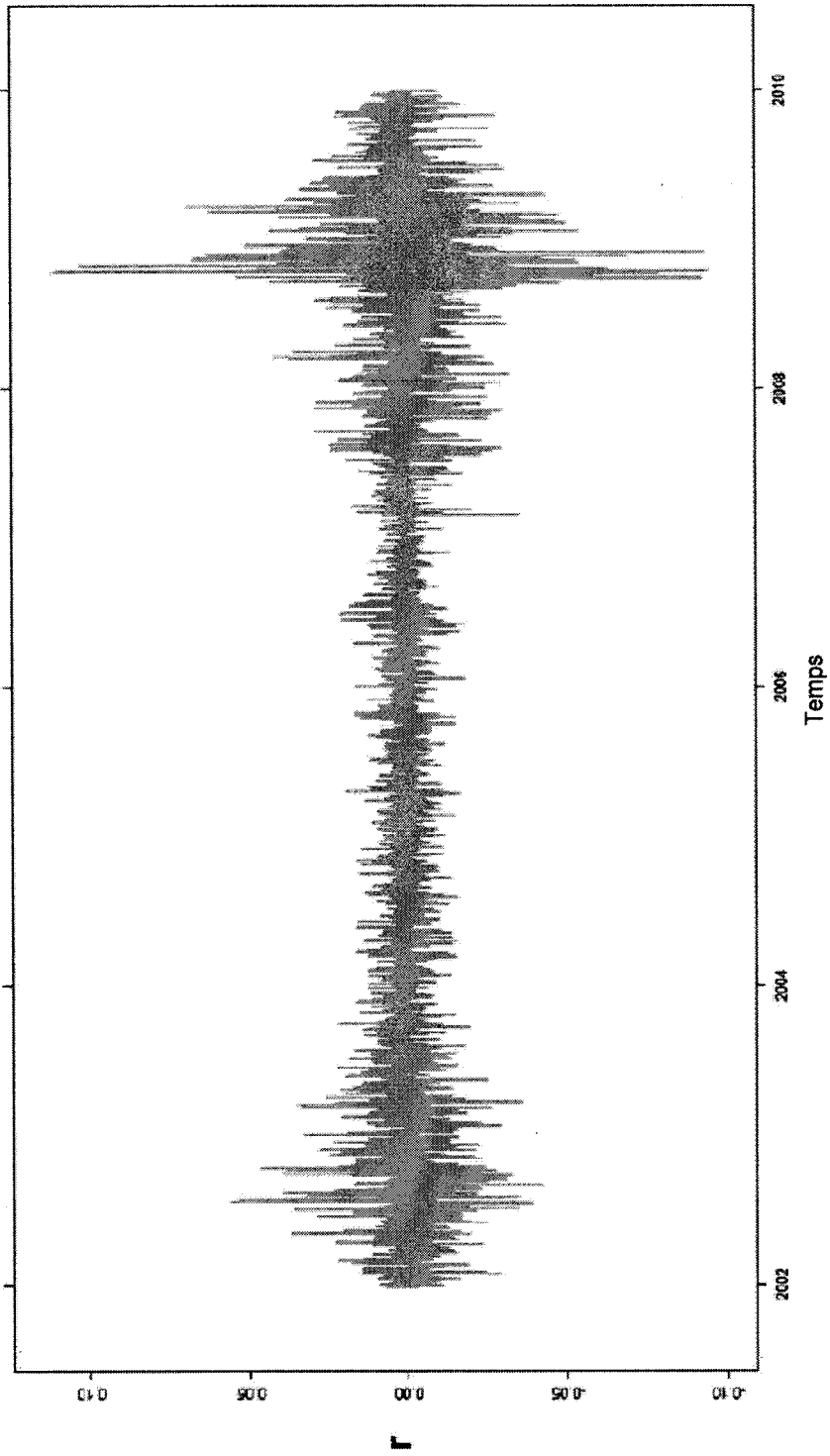


Figure 5a

Stratégies de modèles	MSE	MAE	RMSE	R^A	σ^A	Ratio de Sharpe	nUP	nDown	Rapport de profit
GarchX	0.00020	0.00923	0.01405	17%	22.5%	0.75	1054	942	15.6%
NN-DirStr	0.00019	0.00856	0.01366	32%	22.4%	1.41	1162	834	31.2%
NN-DirStrX	0.00019	0.00872	0.01380	24%	22.4%	1.05	1147	849	22.3%
KNN-DirStr	0.00021	0.00972	0.01467	-15%	22.5%	-0.67	973	1023	-0.12%
KNN-DirStrX	0.00020	0.00932	0.01426	7%	22.5%	0.32	1050	946	0.06%
SVR-DirStr	0.00021	0.00931	0.01441	8%	22.5%	0.35	1058	938	0.07%
SVR-DirStrX	0.00021	0.00940	0.01459	1.5%	22.5%	0.07	1033	965	0.01%
Naïve	0.00044	0.00938	0.02102	-30.5%	22.4%	-1.36	911	1087	-0.23%
Acheter et conserver	-	-	-	-0.3%	22.4%	-0.01	1072	926	-0.003%

Figure 5b

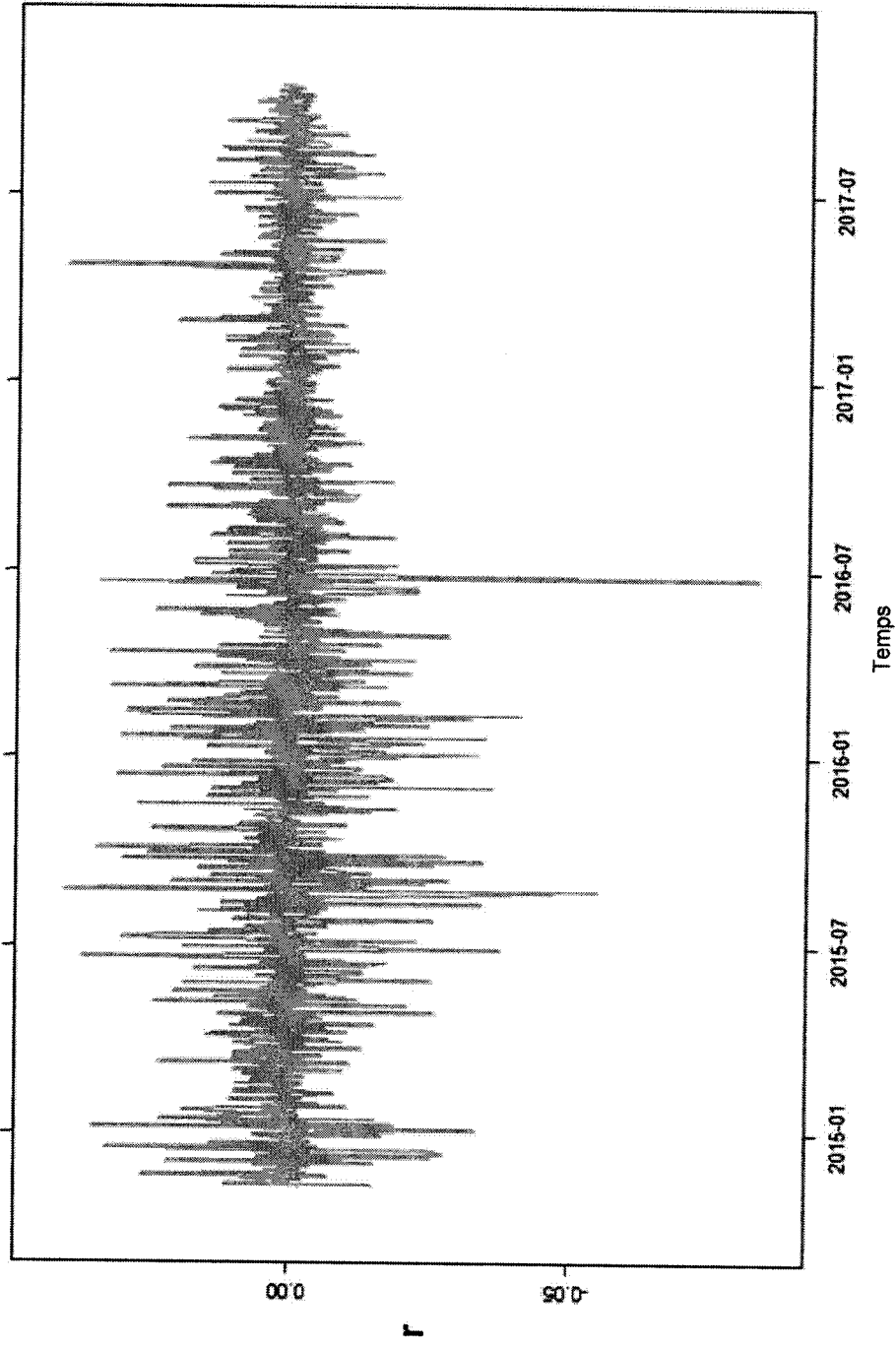


Figure 6a

Stratégies de modèles	MSE	MAE	RMSE	R^A	σ^A	Ratio de Sharpe	nUP	nDown	Rapport de profit
GarchX	0.00015	0.00869	0.01223	10%	19.45%	0.53	386	356	0.1%
NN-DirStr	0.00014	0.00821	0.01193	42%	19.3%	2.15	453	289	46.6%
NN-DirStrX	0.00014	0.00806	0.01180	49.5%	19.21%	2.57	465	277	58.2%
KNN-DirStr	0.00016	0.00916	0.01275	1.6%	19.46%	0.08	361	381	0.02%
KNN-DirStrX	0.00015	0.00876	0.01230	20.2%	19.42%	1.04	384	358	20.3%
SVR-DirStr	0.00015	0.00880	0.01242	-3.4%	19.46%	-0.17	359	383	-0.03%
SVR-DirStrX	0.00016	0.00886	0.0126	-6%	19.46%	-0.3	361	381	-0.05%
Naïve	0.00030	0.00887	0.01743	-3.5%	19.46%	-0.18	359	383	-0.03%
Acheter et conserver	-	-	-	7%	18.58%	0.38	388	354	0.07%

Figure 6b

10/11

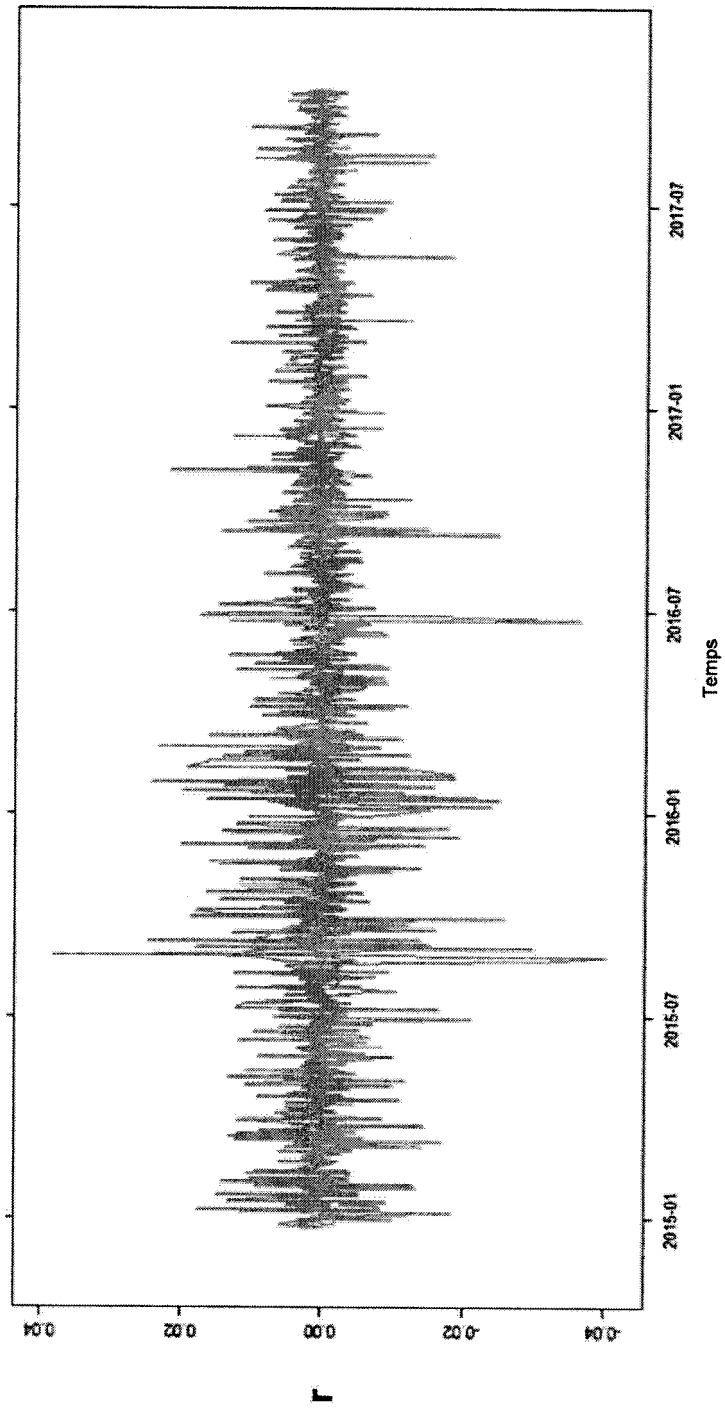


Figure 7a

Stratégies de modèles	MSE	MAE	RMSE	R^A	σ^A	Ratio de Sharpe	nUP	nDown	Rapport de profit
GarchX	0.00066	0.00555	0.00814	0.5%	12.7%	0.04	362	344	0.008%
NN-DirStr	0.00063	0.00543	0.00792	16%	12.7%	1.27	394	312	26%
NN-DirStrX	0.00063	0.00546	0.00795	15%	12.7%	1.20	388	318	24.5%
KNN-DirStr	0.00070	0.00598	0.00839	-5%	12.7%	-0.39	354	352	-0.07%
KNN-DirStrX	0.00067	0.00575	0.00820	-0.9%	12.7%	-0.8	340	366	-0.01%
SVR-DirStr	0.00066	0.00558	0.00812	-10%	12.7%	-0.8	346	360	-0.14%
SVR-DirStrX	0.00066	0.00562	0.00810	-0.02%	12.7%	-0.002	342	365	-0.0003%
Naïve	0.00014	0.0058	0.01205	-9.8%	12.7%	-0.77	319	388	-0.13%
Acheter et conserver	-	-	-	8%	12.3%	0.65	368	339	12.2%

Figure 7b