

Sujet de thèse

Assimilation des dynamiques spatio-temporelles de la turbidité par synergie entre mesures satellitaires de la couleur de l'eau et simulations hydrodynamiques

Laboratoires d'accueil :

- Lab-STICC, CNRS UMR 6285, IMT-Atlantique, Plouzané
- LGO (Laboratoire Géosciences Océan), CNRS UMR 6538, IUEM, Plouzané

Directeur de thèse : Ronan Fablet, Professeur IMT-Atlantique, Plouzané

Co-directeur de thèse : Christophe Delacourt, professeur des Universités, IUEM, Plouzané

Encadrant : Frédéric Jourdin, ingénieur Shom au département SEDIM, Shom, Brest

Directeur du LGO : Marc-André Gutscher, directeur de recherche, IUEM, Plouzané

Résumé

Ce projet de thèse porte sur la reconstruction et le suivi de la turbidité océanique lithogénique¹ par des méthodes d'assimilation « dirigées par les données »² appliquées aux observations satellites de la couleur de l'eau. Le principe général de ces méthodes est de mettre en correspondance des observations satellites avec un ensemble de simulations réalistes des dynamiques hydrodynamiques et sédimentaires pour reconstruire la dynamique de la turbidité. Ces approches apparaissent particulièrement adaptées ici dans la mesure où la complexité des modèles physiques explicites rend leur inversion directe impossible. Ce projet de thèse s'articule autour d'une collaboration entre le Lab-STICC, le Shom et le LGO, avec la complémentarité des expertises associées.

Contacts

Ronan Fablet, lab-STICC, Plouzané

Tel : 02 29 00 12 87

Email : ronan.fablet@imt-atlantique.fr

Frédéric Jourdin, Shom, Brest

Tel : 02 56 31 24 49

Email : jourdin@shom.fr

¹ D'origine principalement minérale, qui provient des sols et des sédiments, par opposition à biogénique.

² Par assimilation « dirigée par les données », nous faisons référence ici à des méthodes d'assimilation [e.g., Lguensat et al., 2017] qui n'exploitent pas de modèles physiques a priori sur la dynamique de l'état d'intérêt. Ces méthodes exploitent des représentations explicites ou implicites de la dynamique de l'état d'intérêt à partir de jeux de données représentatifs de ces dynamiques.

2. Problématique générale

Ce sujet de thèse est principalement basé sur l'exploitation massive des images satellites de la couleur de l'eau. La couleur de l'eau est un indicateur de la production primaire. Observé par satellite, cet indicateur est essentiel à la gestion écosystémique du milieu aquatique. La couleur de l'eau est aussi liée à sa transparence, qui agit sur la disponibilité de la lumière nécessaire à la photosynthèse, et a aussi un effet rétroactif sur le bilan thermique des couches de surface de l'océan via le transfert radiatif de la lumière solaire. La transparence, qui est l'inverse de la turbidité, est aussi le paramètre principal qui détermine les distances de visibilité sous-marines. Dans les eaux superficielles, cette distance est définie par la profondeur de disparition visuelle du disque de Secchi (Lee et al 2015) dont les images satellites de la couleur de l'eau en sont une observation essentielle (e.g. Zheng and DiGiacomo 2017). Les conditions de la visibilité sous-marine intéressent les activités maritimes impliquant des plongeurs ou des systèmes optiques embarqués, dans les domaines de la plaisance, l'ingénierie offshore, la détection sous-marine, les opérations de déminage etc. Cette proposition de thèse s'inscrit ainsi naturellement dans une problématique d'intérêt dual, à la fois scientifique et opérationnel.

Après une période creuse, le nombre de satellites de mesure de la couleur de l'eau mis à disposition de la communauté scientifique croît à nouveau. Avec le lancement du deuxième satellite européen Sentinel-3 (S3) prévu en mars 2018 ce seront alors six capteurs de la couleur de l'eau qui observeront l'océan et fourniront des données aisément accessibles : deux capteurs MODIS, deux VIIRS et deux S3 (sans compter les satellites chinois, indiens, coréens et japonais). En Europe, les capteurs S3, d'une résolution spatiale de 300 mètres, sont désormais complétés par les deux capteurs Sentinels-2 (S2) qui ont une résolution spatiale de 10 mètres pour l'observation du littoral dans le domaine visible. Avec un accès aisé (et gratuit) à une telle masse de données, associé à des capacités importantes de stockage et de traitement de l'information, la couleur de l'eau est susceptible de prendre une part plus importante dans l'océanographie opérationnelle, y compris dans le domaine de la défense.

Le principal obstacle à l'exploitation de l'imagerie satellite de l'océan est la couverture nuageuse. L'objectif de cette thèse est de contribuer aux méthodes d'exploitation massive (domaine du "Big Data") des mesures satellites actuelles et historiques de la couleur de l'eau pour pallier ce problème. La méthodologie proposée est l'assimilation « dirigée par les données ». Il s'agit d'une classe de méthodes émergentes dont uniquement une première évaluation a été réalisée dans le domaine de la couleur de l'eau (Renosh et al 2017). L'équipe TOMS du laboratoire Lab-STICC développe depuis quelques années une compétence dans ce domaine et a introduit l'assimilation analogue (Lguensat et al 2017), qui combine des représentations implicites de type analogue de la dynamique de l'état à assimiler et des algorithmes d'assimilation statistique (e.g., filtres de Kalman). En particulier, nous avons montré la pertinence de la méthode AnKS (Analog Kalman Smoother, Lguensat et al 2017) pour la reconstruction des dynamiques de traceurs à la surface de l'océan (e.g., SST, SSH) à partir d'observations satellites partielles (e.g. Fablet et al 2017, Lguensat et al 2017b). La méthode AnKS permet une évaluation des erreurs a posteriori, qui est une composante essentielle de l'océanographie opérationnelle future (e.g. Le Traon et al 2015). Elle peut également être associée à des techniques de réseaux de neurones (Fablet et al 2017b) et offre un cadre générique pour exploiter d'éventuelles connaissances a priori (e.g., advection-diffusion) (Fablet et al 2018).

Ces travaux de thèse s'effectuent dans le cadre d'une collaboration entre le Shom (Frédéric Jourdin), le Lab-STICC (Ronan Fablet) et le LGO (Christophe Delacourt). Ils contribuent à la convention de collaboration scientifique et pédagogique signée fin 2017 par l'IMT Atlantique et le Shom, pour un partenariat privilégié dans les domaines de la géomatique, le traitement du signal et l'imagerie de l'environnement, ainsi qu'entre le LGO et le département SEDIM du Shom via les chercheurs associés aux deux laboratoires (Thierry Garlan et Frédéric Jourdin).

3. Approche proposée et programme de travail

Le besoin opérationnel associé à cette thèse est la prévision de la turbidité dans la colonne d'eau, avec son niveau d'erreur associé. L'estimation de cette incertitude est un élément de décision essentiel pour un système opérationnel. Elle nécessite l'emploi de méthodes statistiques. Parmi les outils utilisés en océanographie, les formalismes statistiques d'assimilation de données fournissent un cadre générique pour la reconstruction de la dynamique d'un vecteur d'état à partir d'une série d'observations et de l'incertitude associée.

Les techniques classiques d'assimilation de données reposent sur la formulation d'un modèle paramétrique explicite de la dynamique du vecteur d'état (ici, un champ 3D de la turbidité). Elle consiste en général à supposer que l'erreur a priori se situe au niveau d'un ou plusieurs paramètres de contrôle d'un modèle déterministe, lui-même basé sur des équations hydrodynamiques. Les paramètres de contrôle peuvent être par exemple un coefficient de frottement sur le fond, une vitesse de chute des sédiments, ou des conditions aux limites ou conditions initiales. L'assimilation cherche à optimiser les valeurs du ou des paramètres de contrôle pour être en cohérence avec les données d'observation, en tenant compte d'hypothèses faites sur les erreurs a priori dans les données et les paramètres de contrôle. Les équations hydrodynamiques (et plus généralement celles hydrosédimentaires dans notre cas) sont centrales dans le développement de ces méthodes d'assimilation. Sous réserve que les hypothèses de modélisation sont vérifiées, ces méthodes d'assimilation tirent le meilleur parti des observations disponibles, même en quantité restreinte, pour contraindre le modèle. Cependant, elles se révèlent souvent lourdes en termes de développement informatique ou de temps de calcul. Elles sont également peu flexibles par rapport au choix du modèle hydrodynamique et sont mises en défaut en cas d'incohérence entre modèles et données.

Ces différents éléments ont motivé l'exploration d'alternatives moins coûteuses et plus flexibles reposant sur l'analyse d'une série d'observations courantes à l'aune d'un ensemble historique ou simulé représentatif des dynamiques de l'état du système étudié afin d'inférer l'évolution temporelle la plus probable des paramètres recherchés. Dans ces approches dites d'assimilation « dirigée par les données », le jeu de données représentatif des dynamiques de l'état du système étudié constitue une représentation implicite du modèle hydrodynamique explicite considéré par les techniques d'assimilation classique. L'assimilation « dirigée par les données » permet ainsi par exemple de considérer la reconstruction d'un seul paramètre ou variable d'intérêt, alors que le modèle physique fait intervenir l'ensemble des variables d'état du système. C'est ici un avantage puisqu'il permet d'envisager une reconstruction de la turbidité à partir de mesures satellites à la surface de l'océan sans avoir à reconstruire l'ensemble des variables d'état (i.e., densité, courant, production primaire,...) pour des processus d'intérêt tels que la dynamique des panaches de rivière, les épisodes de remise en suspension et ceux de refroidissement.

L'hypothèse sous-jacente principale des méthodes d'assimilation « dirigée par les données » est l'existence d'un jeu de données représentatif de la variabilité du système étudié, et se justifie dans un contexte de type « Big Data ». Cela semble particulièrement adapté à la problématique du suivi de la turbidité, caractérisée par : i) la complexité des modèles physiques explicites réalistes qui rend impossible leur inversion directe à partir des seules observations satellites de surface ; ii) des efforts importants de la communauté scientifique en termes de simulations réalistes des dynamiques hydro-sédimentaires et de production primaire.

Cette thèse explorera différents formalismes d'assimilation « dirigée par les données », notamment ceux basés sur des modèles markoviens à état caché (HMM) (Renosh et al 2017), l'assimilation analogue (Lguensat et al 2017) et les réseaux de neurones (Fablet et al 2017). La pertinence de chaque formalisme sera évaluée vis-à-vis de son interprétabilité en termes de processus géophysiques (e.g., processus d'advection et diffusion) et de la capacité à fournir conjointement une reconstruction de la

turbidité et une estimation de l'incertitude associée. Les modèles retenus et développés seront appliqués à la problématique de l'analyse et du suivi de la turbidité lithogénique ou SPIM (Suspended Particulate Inorganic Matter). Ils exploiteront trois types de données :

Nom	Description des données
SAT	Données des concentrations en SPIM mesurées par les satellites de la couleur de l'eau.
DYN	Données résultats des paramètres dynamiques et de forçages (vagues, courants, décharges fluviales, densité, vent) de modèles purement hydrodynamiques.
SED	Données résultats des concentrations en SPIM de modèles couplés hydrodynamique et sédimentaire (hydrosédimentaires).

Données et Cas d'étude. Nous décrivons plus précisément dans la section suivante les sources de données qui seront considérées dans le cadre de cette thèse pour chaque type de donnée. Les types de données DYN et SED seront exploités en premier pour construire des jeux de données d'apprentissage des dynamiques spatio-temporelles de la turbidité. Ils seront aussi utilisés pour mettre en oeuvre des OSSE (Observing System Simulation Experiment) et disposer de jeux de données d'évaluation pour lesquels une vérité de terrain est disponible. Ce type de jeux de données est particulièrement important dans les phases de calibration et analyse de sensibilité de modèles et algorithmes (par exemple, vis-à-vis du taux de données manquantes dans les mesures satellites dues à la couverture nuageuse). Le jeu de données SAT sera exploité comme jeu de données d'observations des modèles d'assimilation pour des applications sur des données réelles. Nous présentons dans le tableau suivant les différentes expériences numériques qui sont envisagées (en mentionnant explicitement celles de type OSSE) :

Expérience	SAT	DYN	SED
1 (OSSE)			X
2	X		
3	X	X	
4 (OSSE)		X	X
5	X		X
6	X	X	X

NB: dans les *Expériences 1 à 3*, seule la concentration de surface en SPIM sera restituée. Les *Expériences 1 et 4* sont des expériences de type OSSE (Observing System Simulation Experiment).

La mise en oeuvre de ces expériences reposera sur la définition de protocole et critères d'évaluation quantitative et qualitative des reconstructions réalisées. Elles permettront également d'évaluer la sensibilité des méthodes proposées à la qualité du lien entre les modèles hydrodynamiques, hydrosédimentaires et les mesures satellites. Deux zones d'étude privilégiées sont identifiées pour mener des évaluations expérimentales des modèles et algorithmes proposés: le Golfe du Lion (GdL) et le Golfe de Gascogne (GdG). Le choix de ces deux zones est motivé à la fois par leur intérêt thématique et la disponibilité de données de type DYN et SED. La qualité des résultats de validation des modèles hydrosédimentaires disponibles, que ce soit en termes de turbidité de fond, vis-à-vis des mesures in situ, ou en termes de turbidité de surface, vis-à-vis des mesures satellites sera analysée pour définir précisément les expériences mises en oeuvre comme décrits dans la section précédente. Le tableau ci-dessous fournit la liste des produits et simulations identifiés dans le cadre de cette proposition de thèse.

Type	Modèle	Descriptif	Zone	Laboratoire	POC
Base de données satellite	OC5	Serveur du Cersat	GdG	ACRI-HE, Ifremer	Philippe Bryère et Francis Gohin
Base de données satellite	SPM	Kalicôtier	GdL	ACRI-ST	Bertrand Saulquin
Simulation hydro-sédimentaire	Symphonie	Thèse Guillaume Mikolajczak (2018)	GdL	Laboratoire d'Aérodologie	Claude Estournel
Simulation hydrosédimentaire	MARS	Thèse Baptiste Mengual (2016)	GdG	Ifremer	Romaric Verney

Ainsi ces phases expérimentales s'appuieront naturellement sur des collaborations avec des experts thématiques à la fois pour la mise à disposition des données, leur qualification et la définition de critères d'évaluation pertinents, plus précisément :

- P. Bryère (ACRI-HE, Brest) : produits couleur de l'eau de niveau L2/L3/L4
- B. Saulquin (ACRI-ST, Sophia-Antipolis) : expertise en couleur de l'eau
- F. Gohin (Ifremer, Brest) : algorithme couleur de l'eau (OC5) validé en GdG et GdL
- C. Estournel (LA, Toulouse) : modèle hydrosédimentaire Symphonie validé en GdL
- R. Verney (Ifremer, Brest) : modèle hydrosédimentaire MARS validé en GdG
- F. Bourrin (CEFREM, Perpignan) : dynamique sédimentaire en GdL, mesures in situ (navires, mouillages, gliders)

Concernant l'évaluation et la validation des modèles et algorithmes proposées, nous bénéficierons notamment de l'expertise du Shom (F. Jourdin) sur ces thématiques. Nous pourrions considérer des sous-ensembles de données de extraits des jeux listés précédemment, sous-ensembles qui seront uniquement dédiés à cette phase d'évaluation et ne seront pas utilisés par les méthodes d'assimilation. De manière complémentaire, nous pourrions considérer d'autres types de données telles que les mesures in situ suivantes :

Type	Projet	Descriptif	Zone	Laboratoire	POC
Mesures in situ	ANR MATUGLI	Thèse Gaël Many (2016)	GdL	CEFREM (Perpignan)	François Bourrin
Mesures in situ	Projet européen BENTHIS	Thèse Baptiste Mengual (2016)	GdG	Ifremer (Brest)	Romarc Verney

Programme de travail. Le programme de travail envisagé comporte quatre tâches scientifiques principales :

- Tâche 1 : la réalisation d'un état de l'art des travaux dans le périmètre scientifique du sujet de thèse en particulier des méthodes d'assimilation "dirigées par les données";
- Tâche 2 : la comparaison expérimentale ('benchmarking') des méthodes d'assimilation "dirigées par les données" pour un cas de restitution d'une série temporelle du champ de concentrations de SPIM à la surface de l'océan (Expériences 1 à 3, cf. ci-dessus).
- Tâche 3 : le développement et l'évaluation de méthodes des séries temporelles du champ 3D des concentrations en SPIM à partir de séries d'observation satellite (Expériences 4 à 6)
- Tâche 4 : l'amélioration des méthodes proposées et la prise en compte explicite de connaissances physiques (Expériences 1 à 6)

La dernière phase de la thèse sera dévolue à la rédaction du manuscrit et à la préparation de la soutenance.

L'échéancier proposé est le suivant :

	Année 1				Année 2				Année 3			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
Tâche 1 : Bibliographie												
Tâche 2 : analyse de la turbidité en surface												
Article/conférence												
Tâche 3 : analyse de la turbidité en surface et sur la verticale												
Article/conférence												
Tâche 4 : améliorations												
Rédaction												
Soutenance												

3. Références de l'équipe sur le sujet

Ces travaux de thèse s'effectueront entre le Lab-STICC (Ronan Fablet) et le Shom (Frédéric Jourdin) et au sein d'un environnement favorable pour l'étude du paramètre turbidité apporté par l'équipe DYNELI (Direction Christophe Delacourt) du LGO.

Ronan Fablet³, Professeur à l'IMT Atlantique et chercheur au sein de l'équipe TOMS du Lab-STICC, sera le directeur de thèse. Il développe une expertise dans le domaine du traitement du signal et de la télédétection de l'océan, en particulier des méthodes originales d'assimilation "dirigée par les données" [e.g. Lguensat et al 2017 ; Fablet et al 2017] qui ont été récemment appliqué et évalué sur la reconstruction de champs géophysiques à la surface de l'océan (e.g. température, courants géostrophiques). Ces travaux fourniront la base méthodologique des travaux de thèse. Les travaux de thèse bénéficieront de l'environnement scientifique du Lab-STICC et des dynamiques collaboratives nationales et internationales sur ces sujets (e.g., projet OSTST MANATEE, Projet inter-labex SEACS,...).

Christophe Delacourt, Professeur des universités, sera co-directeur de thèse. Il développe une expertise en développements instrumentaux et méthodologiques originaux, principalement en télédétection multi-capteurs et multi-plateformes couplées à des mesures in situ afin d'améliorer la connaissance de la dynamique littorale à différentes échelles et gammes de paramètres physiques et biologiques (e.g. Didier et al 2017, Stumpf et al 2017). L'originalité de son approche est pluridisciplinaire : géomorphologie, géophysique, hydrodynamique, sédimentologie. Elle combine à la fois l'observation et la modélisation, et correspond ainsi à l'approche proposée dans ce sujet de thèse.

Frédéric Jourdin, ingénieur au Shom et chercheur associé au LGO en bio-optique marine, sera co-encadrant de la thèse. Il développe une expertise sur les thématiques de mesure et restitution de la turbidité océanique, notamment pour prédire la visibilité sous-marine (e.g. Fromant et al 2017). Il a récemment initié des travaux exploitant des techniques d'apprentissage automatique pour ces applications (Renosh et al 2017) qui motivent le projet de thèse proposé. Il apportera son expertise thématique aussi bien en termes de spécifications des cas d'étude et protocoles d'évaluations, que de qualification des données et formulations de connaissances a priori pertinentes.

4. Références

Didier D, Bernatchez P., Augereau E., Caulet C., Dumont D, Bismuth E., Cormier L., Floc'h F., Delacourt C., 2017. "LiDAR Validation of a Video-Derived Beachface Topography on a Tidal Flat", *Remote Sens.* 2017 (9), 826; doi:10.3390/rs9080826.

Fablet, R., P.H. Viet, R. Lguensat 2017. Data-driven Models for the Spatio-Temporal Interpolation of satellite-derived SST Fields. *IEEE Transactions on Computational Imaging*. PP. 1-1. doi:10.1109/TCI.2017.2749184.

Fablet R., P. Viet, R. Lguensat, P.-H. Horrein, B. Chapron. Spatio-temporal interpolation of cloudy SST fields using conditional Analog Data Assimilation, 2018. *Remote Sensing*, à paraître.

Fromant, G., F. Floc'h, A. Lebourges-Dhaussy, F. Jourdin, Y. Perrot, N. Le Dantec, and C. Delacourt, 2017: In Situ Quantification of the Suspended Load of Estuarine Aggregates from Multifrequency Acoustic Inversions. *J. Atmos. Oceanic Technol.*, 34, 1625–1643

³ Le cv détaillé de Ronan Fablet est accessible à l'adresse <http://www.imt-atlantique.fr/fr/personne/ronan-fablet>

Lee Z P, Shang S, Hu C, Du K, Weidemann A, W. W, Lin J, Lin G, 2015, “Secchi disk depth: A new theory and mechanistic model for underwater visibility”, *Remote Sensing of Environment*, 169, 139-149, ISSN 0034-4257.

Le Traon PY, D. Antoine , A. Bentamy , H. Bonekamp , L.A. Breivik , B. Chapron , G. Corlett , G. Dibarboure , P. DiGiacomo , C. Donlon , Y. Faugère , J. Font , F. Girard-Ardhuin , F. Gohin , J.A. Johannessen , M. Kamachi , G. Lagerloef , J. Lambin , G. Larnicol , P. Le Borgne , E. Leuliette , E. Lindstrom , M.J. Martin , E. Maturi , L. Miller , L. Mingsen , R. Morrow , N. Reul , M.H. Rio , H. Roquet , R. Santoleri , J. Wilkin, 2015. Use of satellite observations for operational oceanography: recent achievements and future prospects. *Journal of Operational Oceanography*, Vol. 8, Iss. sup1.

Lguensat, R., P. Tandeo, P. Ailliot, M. Pulido, and R. Fablet, 2017, “The Analog Data Assimilation.” *Mon. Wea. Rev.*, 145, 4093–4107.

Many G., 2016, “Impacts of storm and flood events on suspended particulate matter dynamics in the Gulf of Lions. Contributions of gliders to a multi-platform approach”, PhD thesis, Université de Perpignan.

Mikolajczak G., En cours (prévue 2018), “Transfert de la matière particulaire apportée par les fleuves depuis les plateaux continentaux Méditerranéens jusqu’au bassin profond. ”, PhD thesis in preparation, Université de Toulouse.

Mengual B., 2016, “Variabilité spatio-temporelle des flux sédimentaires dans le Golfe de Gascogne : contributions relatives des forçages climatiques et des activités de chalutage”, PhD thesis, Université de Bretagne Occidentale.

Renosh P, Jourdin F, Charantonis A, Yala K, Rivier A, Badran F, Thiria S, Guillou N, Leckler F, Gohin F, Garlan T, 2017, “Construction of Multi-Year Time-Series Profiles of Suspended Particulate Inorganic Matter Concentrations Using Machine Learning Approach”, *Remote Sensing, December 2017*.

Saulquin B, Fablet R, Ailliot P, Mercier G, Doxaran D, Mangin A, Fanton d'Andon O, 2015, “Characterization of Time-Varying Regimes in Remote Sensing Time Series: Application to the Forecasting of Satellite-Derived Suspended Matter Concentrations.” in Selected Topics in *Applied Earth Observations and Remote Sensing, IEEE Journal of* , vol.8, no.1, pp.406-417, Jan. 2015, doi:10.1109/JSTARS.2014.2360239.

Stumpf A., Malet J.P., Delacourt C, 2017. “Correlation of satellite image time-series for the detection and monitoring of slow-moving landslides”, *Remote Sensing of Environment* 189 (2017) 40–55, doi: 10.1016/j.rse.2016.11.007.

Tandéo P, Ailliot P, Ruiz J, Hannart A, Chapron B, Cuzol A, Monbet V, Easton R, Fablet R, 2015. Combining Analog Method and Ensemble Data Assimilation: Application to the Lorenz-63 Chaotic System. *Machine Learning and Data Mining Approaches to Climate Science, Lakshmanan V, Gilleland E, McGovern A, Tingley M, Editors*, Chapter pp 3-12.

Zheng G, DiGiacomo P M, 2017, “Uncertainties and applications of satellite-derived coastal water quality products”, *Progress in Oceanography*, Volume 159, 2017, Pages 45-72, ISSN 0079-6611.