

Modélisation de scène et extraction de source

Sommaire

8.1	Modélisation de scène	2
8.1.1	Présentation de la méthode	2
8.1.2	Projection spatiale du cube intrinsèque	2
8.1.3	Composantes non galactiques de la scène	3
8.1.4	Ajustement de la scene 2D	4
8.1.5	Ajustement chromatique	8
8.2	Extraction des sources	11
8.2.1	Cube de la galaxie hôte	13
8.2.2	Extraction de la Supernova	13
8.3	Classification : SNID	20
8.4	Cas complexes	20
8.4.1	ZTF19acbjnt	21
8.4.2	ZTF19abormno et ZTF20ablhllo	21

Ce chapitre est consacré à la description de la dernière étape du pipeline **HYPERGAL**, la modélisation de scène. Les chapitres précédents ont permis dans un premier temps la construction du cube intrinsèque de la galaxie hôte. Puis nous avons procédé à sa projection dans l'espace spectral de la SEDm à partir de la réponse impulsionnelle spectrale de l'instrument. Enfin, nous avons également construit un modèle de PSF robuste permettant la modélisation de sources ponctuelles.

Dans ce chapitre, nous allons tout d'abord détailler le processus de modélisation de scène, puis nous présenterons les résultats d'extraction des différentes composantes qui la composent. Après avoir illustré ces résultats dans un cas simple, nous montrerons quelques extractions de cas plus complexes obtenues avec **HYPERGAL**.

8.1 Modélisation de scène

8.1.1 Présentation de la méthode

La modélisation de scène implémentée dans **HYPERGAL** va globalement suivre la méthode utilisée pour l'extraction de source ponctuelle, présentée dans le chapitre précédent. Les cubes de données (observations) utilisés sont préalablement calibrés en flux, en utilisant la courbe de sensibilité obtenue à partir de l'étoile standard observée la plus récente. Les rayons cosmiques sont également masqués à l'aide du module **BYECR** (?).

L'idée est de modéliser la scène pour N méta-tranches couvrant un domaine spectral pertinent de la SEDm. En effectuant un ajustement de la scène pour chaque méta-tranche, nous obtiendrons un jeu de $N \times m$ paramètres. Puis, à l'instar de la méthode d'extraction de source ponctuelle, nous procèderons à un ajustement de la chromaticité des différentes composantes de la scène. Cela nous permettra de fixer tous les paramètres de forme et de position. Enfin nous terminerons par un ajustement linéaire des amplitudes des différentes composantes pour toutes les tranches du cube de données.

Cette procédure nécessite dans un premier temps de projeter notre cube intrinsèque dans l'espace de la SEDm. Nous rappelons qu'à l'issue de la détermination de la réponse impulsionale spectrale (LSF), nous avons déjà projeté notre cube intrinsèque dans l'espace spectral de la SEDm. Il nous manque donc la projection dans l'espace spatial.

8.1.2 Projection spatiale du cube intrinsèque

La projection du cube ne se fait pas en une opération, mais en projetant successivement chaque tranche qui le compose.

Il nous faut pour cela prendre en compte la géométrie des spaxels des 2 cubes. Pour les traitements géométriques, nous utilisons le module **SHAPELY**¹ (?), qui nous permet de reconstruire la grille avec les spaxels carrés du cube intrinsèque, et celle avec les spaxels hexagonaux du cube de données SEDm.

Avant de projeter le flux, nous adaptions l'échelle des pixels entre les deux espaces. Nous savons que les pixels des images PS1 ont une taille de $0''.25$ de côté. Afin de connaître précisément le facteur d'échelle à appliquer, nous avons effectué une analyse spatiale sur des observations de la SEDm avec plusieurs sources (> 3) dans le champ de vue. Par comparaison géométrique avec les images PS1 de la même zone du ciel, analogue à une triangulation, nous avons déterminé un rapport d'échelle de 2.230 ± 0.003 entre la taille des spaxels de la SEDm et des images PS1. Nous en avons déduit une taille effective des spaxels hexagonaux de $0''.558/\text{spaxel}$. Il est important de comprendre que cette adaptation d'échelle est purement numérique et ne correspond pas à un ré-échantillonnage. Avant de projeter le flux du cube intrinsèque, nous incluons un modèle d'adaptation de PSF. En effet, les images PS1 ayant un seeing plus petit ($\sim 1''.1$) que celui de la SEDm ($\sim 1''.7$), nous devons prendre en compte cette différence avant le ré-échantillonnage spatial, et inclure une PSF différentielle.

En toute rigueur, il faudrait entraîner un modèle de PSF relatif entre PS1 et la SEDm. Dans ce travail, nous avons supposé que la correction du seeing relatif pouvait être modélisée par une gaussienne 2D asymétrique (présentant une potentielle ellipticité). Le

1. <https://github.com/shapely/shapely>

seeing des images PS1 et de la SEDm n'étant pas fixe, les paramètres $\{\sigma, \mathcal{A}_G, \mathcal{B}_G\}$ de ce modèle seront libres dans notre modélisation de scène.

Après la convolution de la tranche considérée du cube intrinsèque par ce kernel gaussien, nous devons déterminer une position d'ancrage entre la grille hexagonale et la tranche du cube à projeter.

Cette ancre de projection doit être une position du ciel dont nous connaissons la localisation à la fois dans les images PS1, et dans le MLA de la SEDm. Par défaut dans **HYPERGAL**, nous utilisons la position de l'événement transitoire détectée par la caméra ZTF, à partir de laquelle nous avons récupéré les images PS1. Cette position est localisée au centre des images PS1, et donc au centre des tranches du cube intrinsèque. La caméra de guidage de la SEDm (la *Rainbow Camera*) nous fournit également une position approximative de l'objet détecté dans le MLA. Nous alignons ainsi cette position du MLA avec le centre de la tranche du cube considérée avant d'effectuer la projection du flux.

Pour procéder à la projection spatiale du flux dans l'espace spatial de la SEDm, nous utilisons le module **GEOPANDAS**¹ (?). Cet outil nous permet de superposer les deux grilles de polygones décrivant les géométries du cube intrinsèque et du MLA, puis de déterminer les aires de chevauchement entre tous les pixels.

Nous récupérons ainsi pour chaque spaxel du MLA l'intégrale des flux du cube qui le chevauchent, en pondérant par l'aire de superposition. Cette aire de superposition est égale à 1 lorsque qu'un spaxel carré du cube est entièrement contenu dans un spaxel hexagonal du MLA.

Nous présentons dans la Figure 8.1 la projection d'une méta-tranche du cube intrinsèque dans l'espace de la SEDm, convoluée par une gaussienne 2D sans ellipticité d'écart type 1 spaxel (= 0''.5 dans l'espace PS1). L'ancrage est effectué à partir de la position de la supernova ZTF18accrorf dans le MLA estimée par la caméra de guidage.

Nous procédons ainsi à cette projection pour chaque méta-tranche du cube spectral. Tout comme avec les étoiles standards, les cubes de données SEDm sont affectés par les effets d'ADR, et ainsi la position d'ancrage varie en fonction de la longueur d'onde.

Ces paramètres (x_0, y_0) sont donc également des paramètres libres de notre modélisation de scène, et la position renseignée par la *Rainbow Camera* fait office de condition initiale.

À ce stade de la modélisation, toutes les contributions relatives entre PS1 et la SEDm ont été prises en compte. Nous pouvons à présent compléter la scène avec les composantes de fond et de source ponctuelle.

8.1.3 Composantes non galactiques de la scène

8.1.3.1 Composante du fond : ciel et artefacts

Le fond du ciel ayant été retiré dans les images PS1 (?), il nous faut modéliser cette composante dans le cube SEDm. Pour les mêmes raisons que celles évoquées dans le chapitre ?? avec l'extraction des étoiles standards, nous choisissons de modéliser le fond par un polynôme de second degré tel que :

$$\text{Bkgd}(x, y) = b_0 + (b_x \times x) + (b_y \times y) + (b_{xx} \times x^2) + (b_{yy} \times y^2) + (b_{xy} \times xy) \quad (8.1)$$

1. <https://geopandas.org/>

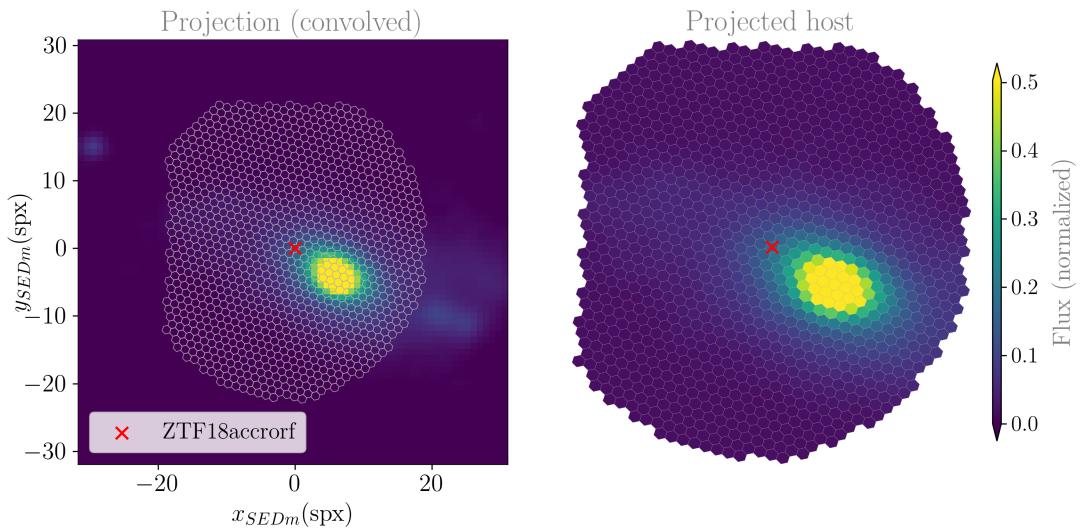


FIGURE 8.1 – Projection de la galaxie hôte dans le MLA pour une méta-tranche du cube intrinsèque. Pour cet exemple, nous avons convolué la tranche par une gaussienne 2D sans ellipticité d'écart type 1 spaxel. La croix rouge indique la position de la supernova ZTF18accorf estimée par la *Rainbow Camera* dans le MLA, qui sert d'ancrage à la projection d'un espace spatial à l'autre.

avec x et y les coordonnées dans le MLA. La constante b_0 est donc utilisée pour modéliser le fond constant du ciel, tandis que les autres paramètres sont là pour corriger les artefacts présents dans le cube de données.

8.1.3.2 Composante de la supernova

L'autre composante de la scène n'est autre que la supernova, une source ponctuelle entièrement caractérisée par la PSF de la SEDm. Nous utilisons donc bien évidemment le profil radial constraint établit au chapitre ??, et défini par l'équation ??.

La position de la supernova à modéliser dans le MLA est supposée confondue avec la position d'ancrage utilisée lors de la projection du cube intrinsèque. Cette approximation signifie que nous considérons la position de détection dans le ciel par la caméra ZTF suffisamment précise et ne nécessitant pas de donner de liberté à la position relative entre la galaxie et l'objet détecté.

8.1.4 Ajustement de la scène 2D

Toutes les composantes de la scène ayant été décrites, nous pouvons passer à l'ajustement de la scène 2D, en considérant les méta-tranches indépendamment les unes des autres.

Nous avons choisi par défaut dans `HYPERGAL` de considérer 6 méta-tranches dans l'intervalle spectral $\lambda \in [5000, 8500]$ Å, où le signal sur bruit est le plus élevé (?). Ce choix d'intervalle est également motivé par la précision spectrale de la calibration en flux de la Figure ?? du chapitre précédent, optimale sur ce domaine spectral. Bien que notre modèle de fond ait été conçu pour prévenir de potentiels artefacts structurés, ceux-ci deviennent parfois trop intenses en dehors de ce domaine spectral. En se restreignant à cet intervalle, nous réduisons le risque de valeurs d'ajustement aberrantes pouvant compliquer

par la suite l'ajustement chromatique. De plus, la majorité des supernovae observées par la SEDm sont des SNIea ($\sim 75\%$), dont la magnitude diminue fortement vers le rouge, notamment au delà de 8000-8500Å dans leur référentiel propre¹. Le contraste entre une SNIa et sa galaxie hôte dans une méta-tranche au delà de ces valeurs est donc fortement réduit, ce qui permet difficilement de contraindre les paramètres de forme de la PSF. La Figure 8.2 illustre bien le type d'artefacts auxquels nous faisons référence, notamment sur les bords des cubes dans le bleu (fond structuré non physique et intense), et des franges dans le rouge, probablement de Moiré (phénomène d'échantillonnage lors de l'extraction du cube des images CCD).

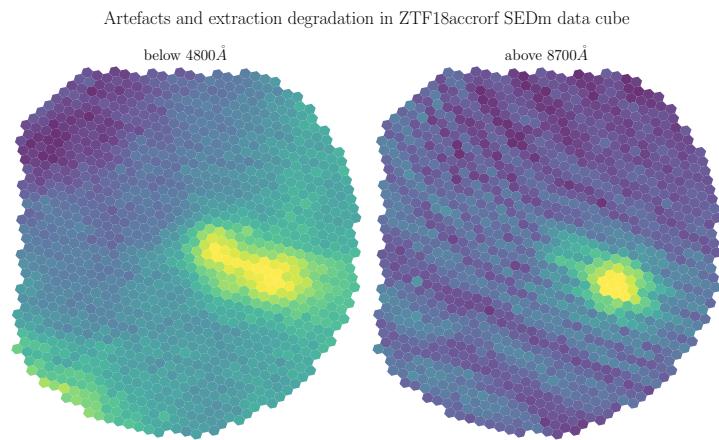


FIGURE 8.2 – Exemple d'artefacts dans les cubes de données SEDm pour ZTF18accorf. Nous montrons à gauche toutes les tranches en dessous de 4800Å du cube de données empilées, et à droite, toutes celles au dessus de 8700Å. Nous pouvons clairement voir dans le bleu la présence d'un signal de fond structuré et intense, distordant complètement les objets dans le champ de vue, et causant la présence d'artefacts sur les bords du cube. Dans le rouge la forme des sources est également altérée, mais nous voyons surtout un phénomène de franges dans les données.

Nous présentons dans la Table 8.1 la liste des paramètres libres de la modélisation de scène pour une méta-tranche.

L'ajustement se fait par minimisation de χ^2 , défini comme :

$$\chi^2 = \sum_{sp} \left(\frac{y_{sp} - \tilde{y}_{sp}}{\sigma_{sp}} \right)^2 \quad (8.2)$$

où y_{sp} et σ_{sp} sont respectivement le flux et la racine de la variance dans un spaxel sp de la méta-tranche du cube SEDm, et \tilde{y}_{sp} le flux modélisé dans ce même pixel. Nous effectuons la minimisation avec le module `IMINUIT`² (??).

La procédure de construction de la scène pour une méta-tranche et pour chaque pas de minimisation du χ^2 se fait de la façon suivante :

¹. L'impact dans le référentiel d'observation est minimum étant donné que la profondeur en redshift de la SEDm se limite à $z \leq 0.1$.

². <https://iminuit.readthedocs.io/en/stable/>

TABLE 8.1 – Paramètres de modélisation de scène incluant toutes les composantes pour une méta-tranche dans **HYPERGAL**.

Paramètre	Symbol
Géométrie	
Position d'ancrage	x_0, y_0
Galaxie hôte	
PSF relative SEDm/PS1	σ
Ellipticité (PSF relative)	$\mathcal{A}_G, \mathcal{B}_G$
Amplitude	G
Fond	
Artefacts	$b_{xx}, b_{yy}, b_{xy}, b_x, b_y$
Ciel	b_0
Source ponctuelle (SN)	
PSF SEDm	α, η
Ellipticité (PSF)	\mathcal{A}, \mathcal{B}
Amplitude	I

Note. Les paramètres d'amplitudes et de fond sont des paramètres de nuisance dans la modélisation des méta-tranches.

- (a) Convolution de la méta-tranche du cube intrinsèque par la PSF différentielle : kernel gaussien 2D asymétrique décrit par les paramètres de forme $\{\sigma, \mathcal{A}_G, \mathcal{B}_G\}$, avec un paramètre d'amplitude G ;
- (b) Projection dans l'espace spatial de la SEDm suivant à partir de la position d'ancrage, décrit par les coordonnées $\{x_0, y_0\}$;
- (c) Ajout du fond structuré (ciel + artefacts), dont les paramètres sont les b_i du modèle quadratique de l'équation 8.1 ;
- (d) Ajout de la source ponctuelle (PSF + amplitude), décrite par les paramètres de forme $\{\alpha, \eta, \mathcal{A}, \mathcal{B}\}$ et l'amplitude I ;
- (e) Détermination du χ^2 , puis réitération des étapes précédentes ;

Nous présentons dans la Figure 8.3 l'ajustement des 6 méta-tranches par **HYPERGAL** pour ZTF18accorf. Pour chaque longueur d'onde, nous montrons la scène observée, la scène ajustée, et le résidu avec le RMS spatial relatif associé. Dans le cas de cette observation, le RMS spatial varie de 1.8% à 3.9%. Nous pouvons également remarquer l'augmentation de RMS pour la méta-tranche la plus *bleue* centrée sur 5285Å, probablement à cause d'un fond de plus en plus structuré comme illustré dans la Figure 8.2. La méta-tranche la plus *rouge* à 8200Å commence également à présenter des artefacts structurés, vraisemblablement des franges de Moiré (phénomène d'interférences).

La Figure 8.4 quant à elle présente la matrice de corrélation entre les paramètres libres de la scène pour la méta-tranche à $\lambda = 6461\text{\AA}$. L'idéal serait bien entendu que tous nos paramètres soient indépendants. Il s'avère que c'est le cas pour la majorité d'entre eux, mais nous relevons cependant quelques corrélations significatives. La première provient des paramètres du fond structuré, et notamment entre les composantes b_{xx} et b_0 , et b_{yy} et b_0 . Pour réduire ces corrélations, nous pourrions par exemple entraîner ce modèle de

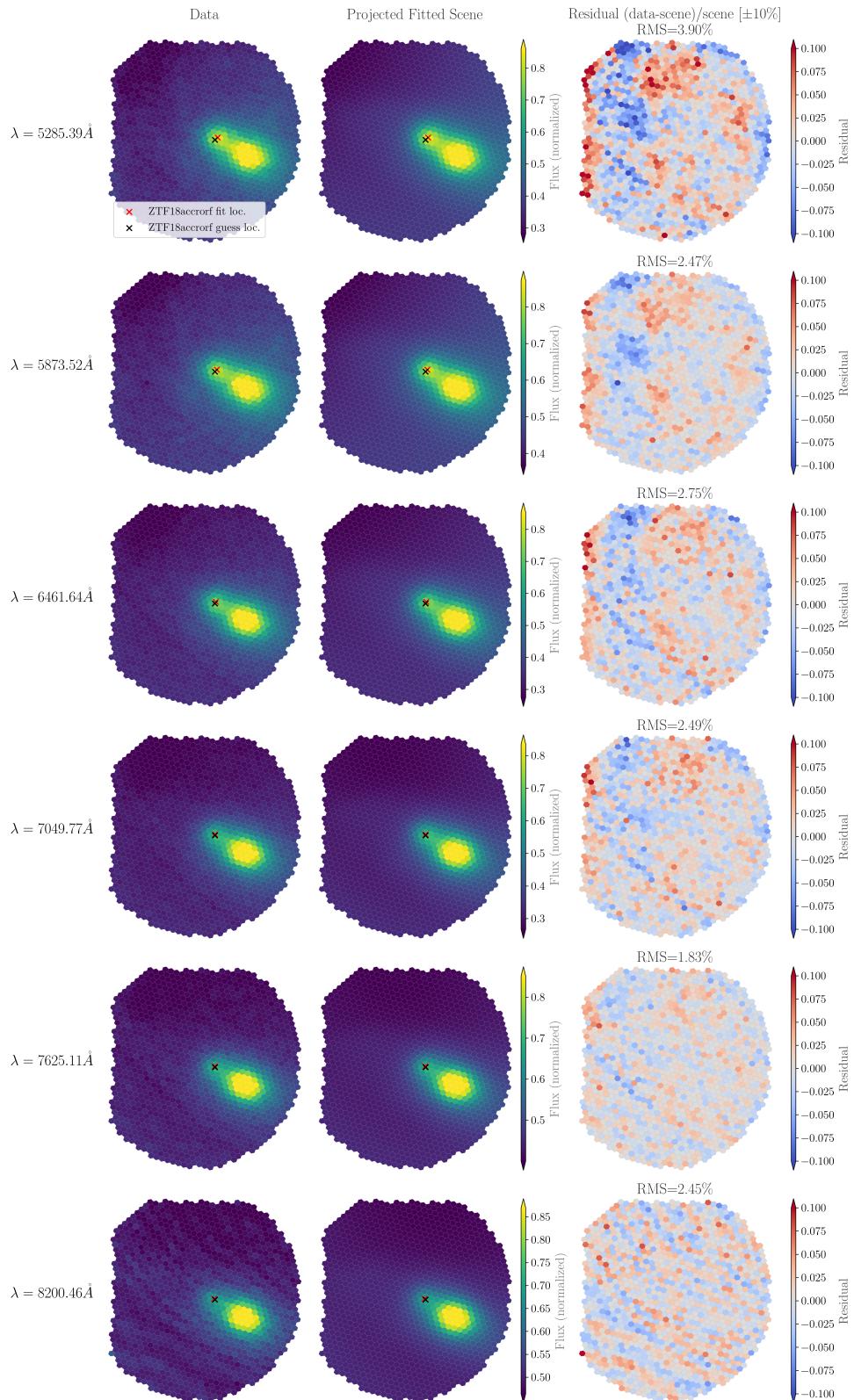


FIGURE 8.3 – Ajustement des méta-tranches pour la modélisation de scène de ZTF18accorf. *De haut en bas* sont représentées les méta-tranches modélisées du bleu vers le rouge. Pour chaque ligne *de gauche à droite* : La méta-tranche modélisée par HYPERGAL, la méta-tranche du cube de données SEDm, et le résidu pondéré par le modèle. Nous indiquons pour chaque longueur d'onde le RMS spatial de l'ajustement, allant de 1.8% à 3.9%. Les croix noires et rouges indiquent respectivement la position d'ancrage initiale (caméra de guidage), et la position ajustée par HYPERGAL.

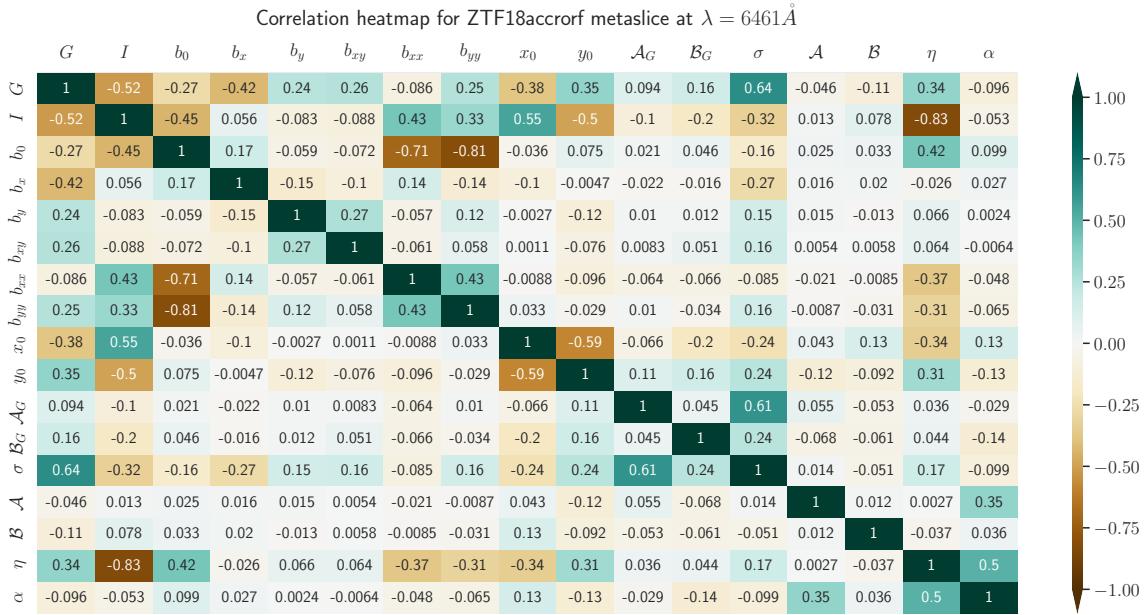


FIGURE 8.4 – Matrice de corrélation des paramètres d’ajustement de scène de ZTF18accrorf pour la méta-tranche à $\lambda = 6461\text{\AA}$.

fond sur de nombreuses observations et procéder à une analyse en composantes principales (PCA), et donc modéliser ce fond structuré dans une autre base.

L’autre corrélation importante survient entre η (poids entre la gaussienne et la Moffat du profil de PSF de la source ponctuelle) et l’amplitude I de la source ponctuelle. Nous pensons que lorsque les ailes du profil radial sont peu intenses devant le signal de fond, alors le paramètre η possèderait trop de liberté satisfaisant le profil du coeur de la source ponctuelle. Le coefficient de normalisation du profil de PSF dépendant de η , cela pourrait expliquer cette dégénérescence avec l’amplitude. Ce phénomène est en cours d’investigation, et il est en effet possible que fixer ce paramètre soit nécessaire pour lever la dégénérescence entre le coeur du profil (la gaussienne) et ses ailes (la Moffat).

8.1.5 Ajustement chromatique

L’ajustement de chaque méta-tranche nous donne ainsi un jeu de $N \times m$ paramètres sur le domaine spectral [5000,8500]Å. Nous modélisons la chromaticité de la PSF de la source ponctuelle de la même manière qu’avec les étoiles standards : une loi de puissance pour α (rayon de la Moffat), et une constante pour η (poids gaussienne/Moffat), \mathcal{A} et \mathcal{B} (décrivant simultanément l’ellipticité et l’orientation).

Nous montrons dans la Figure 8.5 la modélisation chromatique de ces 4 paramètres.

Nous utilisons également pour l’ellipticité et l’orientation de la PSF différentielle (SEDm/PS1) une modélisation achromatique (par une constante).

La chromaticité du rayon de la gaussienne de cette PSF différentielle est modélisée par une loi de puissance, de la même façon que le paramètre de forme principal de la source ponctuelle :

$$\sigma(\lambda) = \sigma_{ref} \left(\frac{\lambda}{\lambda_{ref}} \right)^{\rho_g} \quad (8.3)$$

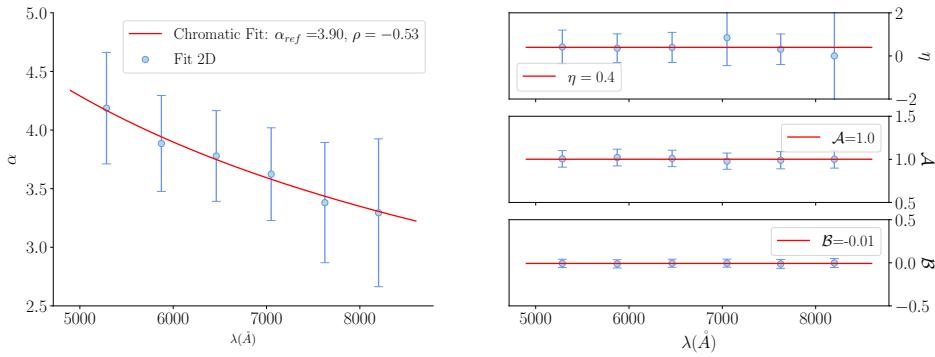


FIGURE 8.5 – Ajustement de la chromatique des paramètres de forme de la PSF pour la supernova ZTF18accrorf, à partir des 6 méta-tranches. À gauche l'ajustement du paramètre α avec une loi de puissance. À droite de haut en bas, l'ajustement par une constante de η (poids entre la gaussienne et la Moffat) et des paramètres A et B .

De la même façon que pour la source ponctuelle, nous montrons dans la Figure 8.6 la modélisation chromatique des paramètres décrivant la PSF relative par laquelle nous avons convolué les méta-tranches du cube intrinsèque.

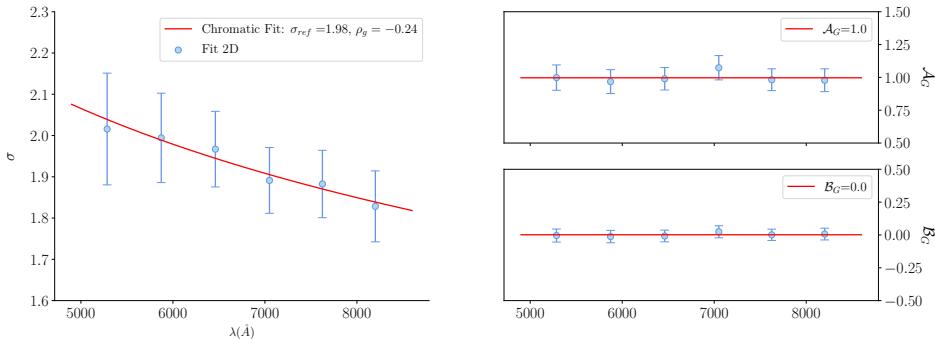


FIGURE 8.6 – Ajustement de la chromatique des paramètres de forme de la PSF différentielle SEDm/PS1 pour la galaxie hôte de la supernova ZTF18accrorf, à partir des 6 méta-tranches. À gauche l'ajustement du paramètre σ avec une loi de puissance. À droite de haut en bas, les paramètres A_G et B_G .

Enfin, à partir de l'ajustement des positions d'ancre $\{x_0, y_0\}(\lambda)$ lors de la projection pour chaque méta-tranche, nous modélisons les paramètres de la réfraction atmosphérique différentielle pour cette observation. Cet ajustement est illustré dans la Figure 8.7.

Si pour une raison quelconque l'ajustement d'une méta-tranche n'a pas convergé, celle-ci est ignorée lors de l'ajustement chromatique. D'autre part, nous limitons l'impact de potentielles valeurs aberrantes pour l'ajustement chromatique en utilisant une fonction de perte de Huber (?). Plus exactement, nous utilisons la fonction de perte *pseudo-Huber* d'ordre 1 (appelée *soft-l1*), qui est une approximation lisse de la fonction originale, définie elle par morceau.

La fonction de perte standard pour un χ^2 est quadratique, c'est à dire que la perte évolue quadratiquement pour un résidu a ($L(a) = a^2$). C'est ce qui la rend très sensible aux valeurs aberrantes, pouvant rapidement dominer la perte totale. La fonction pseudo-Huber

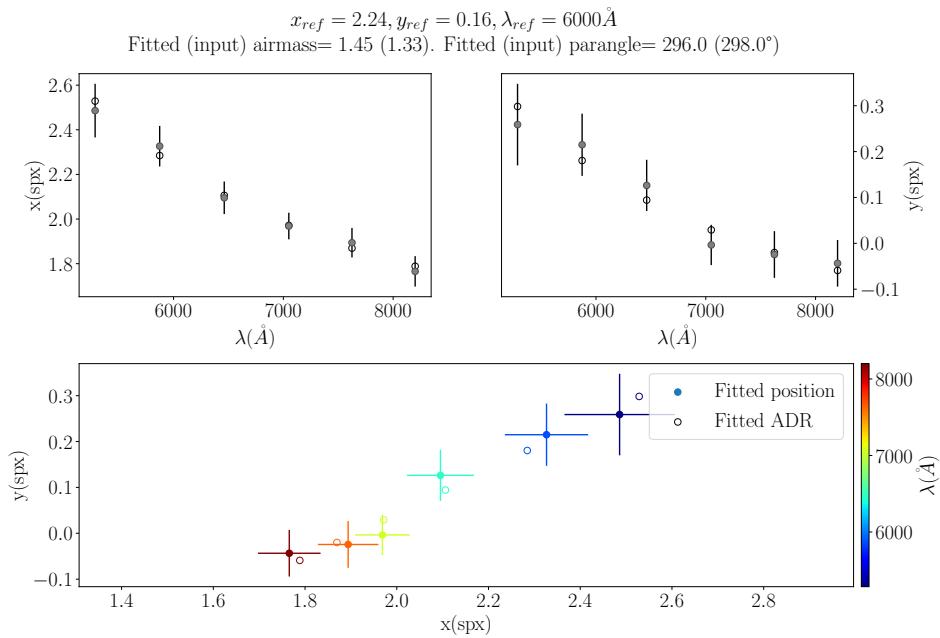


FIGURE 8.7 – Modélisation de la réfraction atmosphérique différentielle pour ZTF18accorf. Nous montrons ici l’ajustement de la position d’ancrage (x_0, y_0) qui représente le centroïde de la supernova par le modèle d’ADR présenté dans le chapitre ???. Les deux graphes du haut représentent l’ajustement des deux coordonnées en fonction de la longueur d’onde. Le graphe du bas illustre l’effet de la réfraction atmosphérique.

à l’ordre 1 est définie comme :

$$H_1(a) = 2 \left(\sqrt{1 + a^2} - 1 \right) \quad (8.4)$$

pour un résidu a . Lorsque le résidu est faible, cette fonction de perte évolue quadratiquement en a^2 , puis évolue linéairement avec une pente de 1 aux grands résidus. Pour un jeu de données x_i associés aux erreurs σ_i , et les points de modélisation m_i , le χ^2 se réécrit alors :

$$\chi_{H_1}^2 = \sum_i H_1 \left(\frac{x_i - m_i}{\sigma_i} \right) \quad (8.5)$$

La Figure 8.8 montre la forme d’un potentiel standard (fonction de perte quadratique), en comparaison avec une fonction de perte pseudo-Huber d’ordre 1. On voit que l’utilisation de ce potentiel permet d’affaiblir le poids des valeurs aberrantes, ce qui permet de ne pas faire dévier excessivement l’ajustement du modèle.

Estimation des valeurs initiales pour la modélisation de scène des méta-tranches.

Toute la procédure expliquée précédemment, de l’ajustement 2D des méta-tranches à l’ajustement chromatique, est également réalisée directement entre les images PS1 et les méta-tranches à transmission équivalentes du cube SEDm. Nous effectuons cette étape préliminaire simultanément avec la construction du cube intrinsèque afin d’optimiser le temps de calcul et les ressources numériques utilisées. Ces ajustements permettent d’obtenir un jeu de paramètres initial pour l’ajustement de scène principal.

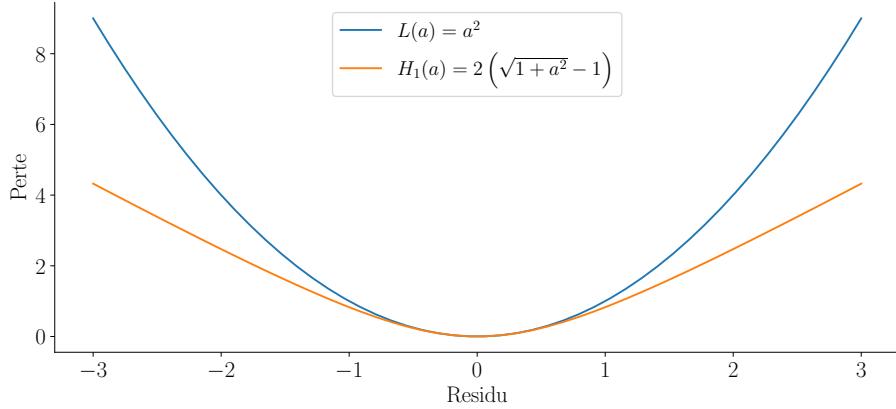


FIGURE 8.8 – Mise en évidence du poids plus faible accordé aux valeurs aberrantes avec la fonction de perte *pseudo-Huber* d’ordre 1.

Une fois l’ajustement de la chromaticité des paramètres de la scène effectué, nous fixons tous les paramètres et procédons à l’ajustement de toutes les tranches du cube, en laissant libre les paramètres d’amplitude uniquement. Le coefficient d’amplitude correctif de la galaxie G et également laissé libre pour toutes les tranches. Nous avons remarqué que cela permet de prévenir les potentiels sur/sous-estimations de l’intensité des raies modélisées dans le cube intrinsèque, d’éventuels fluctuations causées par une calibration en flux de mauvaise qualité ou encore des résidus telluriques dans le cube SEDm. Nous reviendrons sur ce point à la fin de cette section lors de l’extraction des sources pour illustrer nos propos.

La Figure 8.9 montre le résultat final de la modélisation de scène effectuée avec **HYPERGAL** pour la supernova ZTF18accrorf. Nous y présentons l’image 2D du cube de données SEDm et du cube modélisé intégré entre 5000Å et 8500Å. Afin de contrôler la qualité de l’ajustement, nous montrons également le pull spectral et le RMS spectral pour chaque spaxel.

Le RMS spectral est calculé comme dans l’équation ???. Le pull spectral est quant à lui calculé après intégration du spectre pour un spaxel donné de la façon suivante :

$$p_{spx} = \frac{\sum_{\lambda} (y_{\lambda} - \tilde{y}_{\lambda})}{\sqrt{\sum_{\lambda} \sigma_{\lambda}^2}} \quad (8.6)$$

avec \tilde{y}_{λ} la prédiction du modèle, y_{λ} la donnée dans le cube SEDm et σ_{λ} l’erreur sur y_{λ} à la tranche de longueur d’onde λ .

8.2 Extraction des sources

Une fois la modélisation de scène complétée, résultant en un cube 3D dans l’espace des observations de la SEDm, nous sommes en mesure de séparer chacune des composantes : le fond, la galaxie hôte et la source ponctuelle.

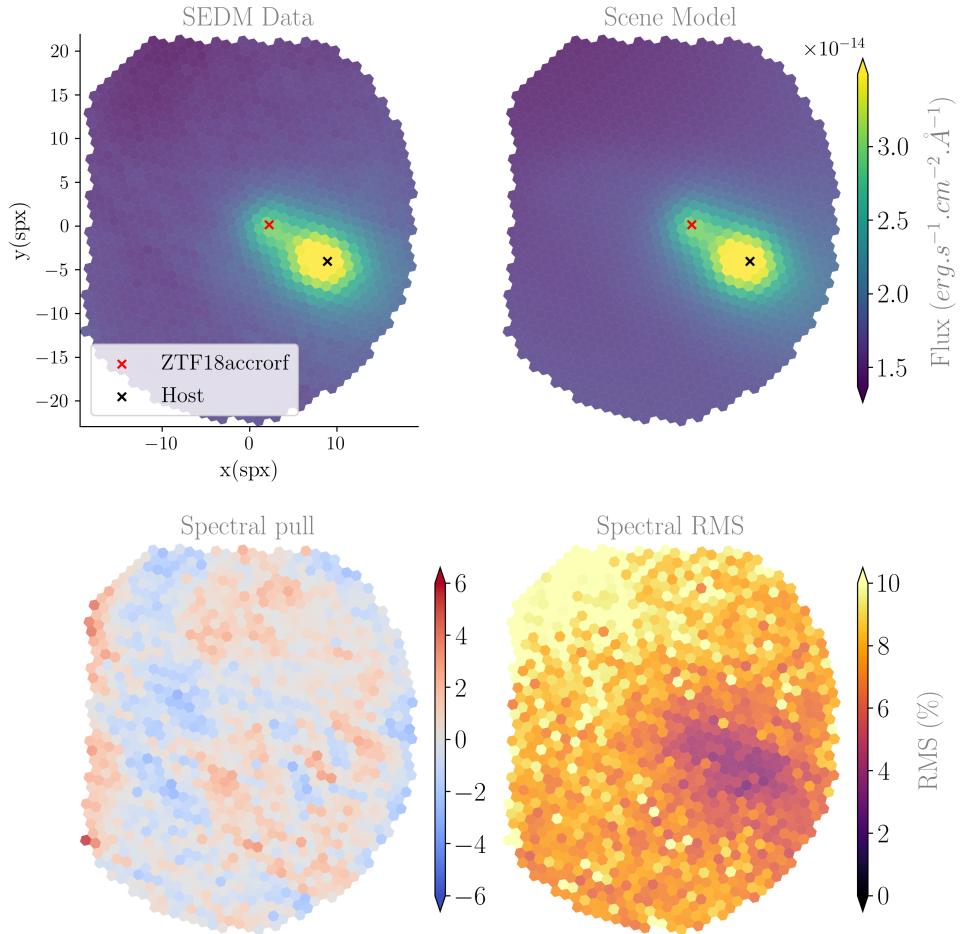


FIGURE 8.9 – Modélisation de scène complète pour ZTF18accorrf. *En haut* nous montrons : le cube de données SEDm intégré entre 5000 et 8500 Å, puis le cube modélisé par HYPERGAL intégré également sur le même domaine spectral. La croix rouge indique la position ajustée de ZTF18accorrf à 7000 Å, et la croix noire la localisation de la galaxie hôte. *En bas* nous montrons : le pull spectral tel que défini dans l'équation 8.6, et le RMS spectral en %. Le pull nous permet de contrôler la présence éventuelle de structures dans le résidu, ce qui n'est pas le cas ici. Le RMS quant à lui nous indique une précision de l'ordre de 4% sur le domaine spectral considéré au niveau des sources présentes dans le champ de vue de la SEDm. Les fluctuations du fond induisent un RMS spectral de l'ordre de 6-7%, et nous pouvons clairement voir les conséquences de l'artefact en haut à gauche du cube de données.

8.2.1 Cube de la galaxie hôte

Afin d'extraire la galaxie hôte du cube de données, nous l'isolons en soustrayant les modèles de fond et de la source ponctuelle au cube SEDm.

Une galaxie n'étant pas une source ponctuelle, nous devons définir une ouverture afin d'en extraire le spectre intégré.

Nous utilisons pour cela l'outil **SEP**¹ (?) (implémentation python de **SEXTRACTOR** ?), en définissant une ellipse d'ouverture dans les images PS1 que nous projetons ensuite dans le MLA de la SEDm à l'aide des solutions WCS des deux espaces. Nous négligeons les effets d'ADR dans cette procédure, étant donné qu'ils induisent rarement un déplacement de plus d'un spaxel dans le champ de vue. Nous illustrons dans la Figure 8.10 l'extraction du spectre intégré de la galaxie hôte de ZTF18accorf, en considérant les spaxels mis en évidences sur l'image du cube intégré. Ce que nous montrons ici n'est pas le spectre de la galaxie modélisée, mais bien celui de la galaxie dans le cube de données de la SEDm, auquel nous avons retiré les modèles de fond et de la source ponctuelle.

De la même façon que la modélisation hyperspectrale de la galaxie (et du fond) nous permet de lever la contamination de la supernova, la modélisation de la supernova nous permet également de réduire la contamination de la galaxie et de l'isoler dans les cubes d'observation.

Connaissant également a priori le redshift utilisé pour la modélisation de la galaxie, nous indiquons la position déduite de quelques raies d'absorption et d'émission^{2, 3} dans l'air (?), afin de visualiser la cohérence entre le spectre isolé dans les données et le redshift utilisé. Dans le cas de cette galaxie, nous pouvons par exemple voir la concordance entre le redshift $z = 0.042$ et la position de la raie H α .

8.2.2 Extraction de la Supernova

De la même façon qu'avec la galaxie hôte, nous pouvons vérifier la séparation de la supernova dans le cube de données en y soustrayant le modèle de fond et de la galaxie. La Figure 8.11 illustre ainsi le cube SEDm de l'observation auquel nous avons retiré les deux autres composantes, et nous pouvons voir à quel point la supernova est bien définie sans structure résiduelle apparente. Nous pouvons par exemple définir une ouverture arbitraire (ici circulaire de 8 spaxels $\sim 4''5$ de rayon) centrée sur la position de la source ponctuelle et porter un regard sur le pull et le RMS spectral dans cette ouverture. Pour cette observation nous pouvons par exemple voir un RMS spectral de l'ordre de 3-4% au niveau de la position de la supernova dans le cube, et aucune structure résiduelle apparente (en provenance d'une mauvaise modélisation galactique par exemple).

Nous pouvons également vérifier la robustesse du modèle de PSF que nous avons défini et contraint dans le chapitre précédent en superposant le modèle ajusté de profil radial aux données. Afin de permettre également un contrôle de la qualité de l'ajustement du fond de ciel, et la présence éventuelle de structure résiduelle, nous visualisons ce profil radial après soustraction du fond ajusté et du modèle hyperspectral de la galaxie hôte. Nous montrons ainsi dans la Figure 8.12 le profil radial ajusté pour une des métatranches. Nous pouvons voir que le profil radial tend bien vers 0, ce qui indique une bonne estimation de cette

1. <http://github.com/kbarbary/sep>

2. <http://astronomy.nmsu.edu/drewski/tableofemissionlines.html>

3. <http://classic.sdss.org/dr6/algorithms/linetable.html>

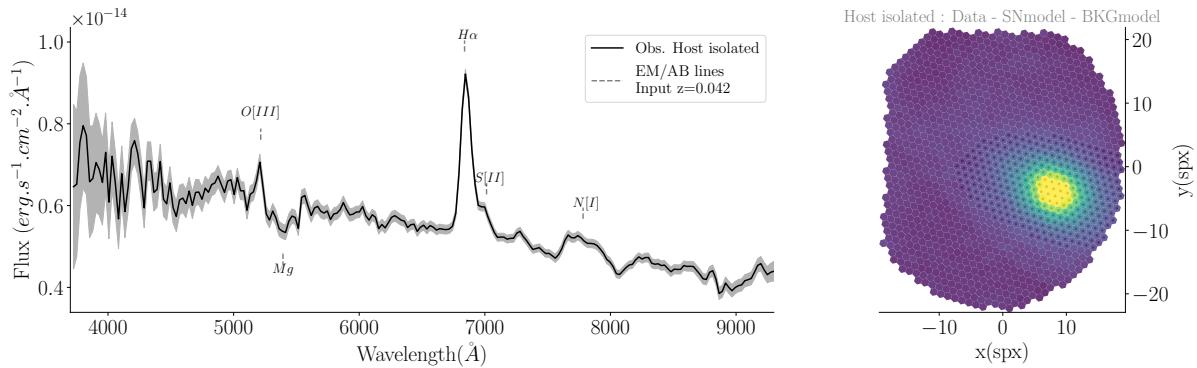


FIGURE 8.10 – Extraction du spectre intégré de la galaxie hôte de ZTF18accrorf. *À droite* nous montrons le cube de données de la SEDm, auquel nous avons soustrait les modèles de fond et de la source ponctuelle afin d’isoler la galaxie hôte. Les spaxels sélectionnés par les losanges noirs indiquent ceux appartenant à l’ouverture utilisée pour l’extraction du spectre intégré. *À gauche* nous montrons donc le spectre extrait et l’erreur associée contenus dans cette ouverture. Tout le domaine spectral de la SEDm est affiché, bien que nous avons déjà discuter de la faible fiabilité de l’extraction aux extrémités du domaine couvert par la SEDm due aux artefacts présents dans les cubes. Nous montrons également la position théorique de quelques raies d’émission et d’absorption sachant le redshift utilisé ($z = 0.042$), et nous pouvons voir ici la bonne cohérence avec le spectre intégré extrait.

composante. Par ailleurs, nous n’observons pas de valeur aberrante dans la supernova isolée, ce qui à son tour traduit une soustraction galactique sans résidu notable.

L’ajustement de l’amplitude de la PSF de la supernova à chaque tranche nous permet ainsi d’en extraire le spectre, de la même façon qu’avec les étoiles standards. Nous montrons enfin dans la Figure 8.13 le spectre extrait de ZTF18accrorf.

Dans la Figure 8.14 nous montrons la superposition du spectre des 3 composantes (fond, galaxie, supernova), ainsi que l’ajustement du coefficient de correction G du cube intrinsèque. L’objectif principal de cette visualisation et de vérifier une potentielle contamination entre les spectres, par exemple des raies d’émission de la galaxie, ce qui ne semble pas être le cas ici. Nous pouvons par ailleurs voir l’évolution chromatique du coefficient G , qui semble corriger dans le cube intrinsèque un excédant d’intensité de la raie d’émission O[III] vers 5200Å, et un déficit d’intensité de la raie H α . Bien que nous n’ayons pas poussé l’analyse de cet effet, nous avons choisi pour le moment de laisser ce coefficient G libre lors de l’ajustement linéaire des amplitudes par tranche spectrale, afin d’avoir cette liberté de correction. Le modèle de ciel présent dans cette figure correspond au coefficient de degré 0 du modèle de fond b_0 (équation 8.1).

Nous présentons finalement le schéma complet de toute la procédure de modélisation de scène avec HYPERGAL. Un script d’automatisation de tout ce processus est également disponible dans le code du pipeline, prenant en entrée le nom d’une cible observée avec ZTF (par exemple ZTF18accrorf), et/ou le chemin d’accès au cube de données. Les informations relatives à l’événement transitoire étudié (redshift, RA/DEC) sont automatiquement récupérées sur le serveur Fritz¹ (????), mais peuvent également être imposées en argument. HYPERGAle est conçu pour être suffisamment flexible et adaptable à n’importe quel IFU.

¹. <https://fritz.science/>

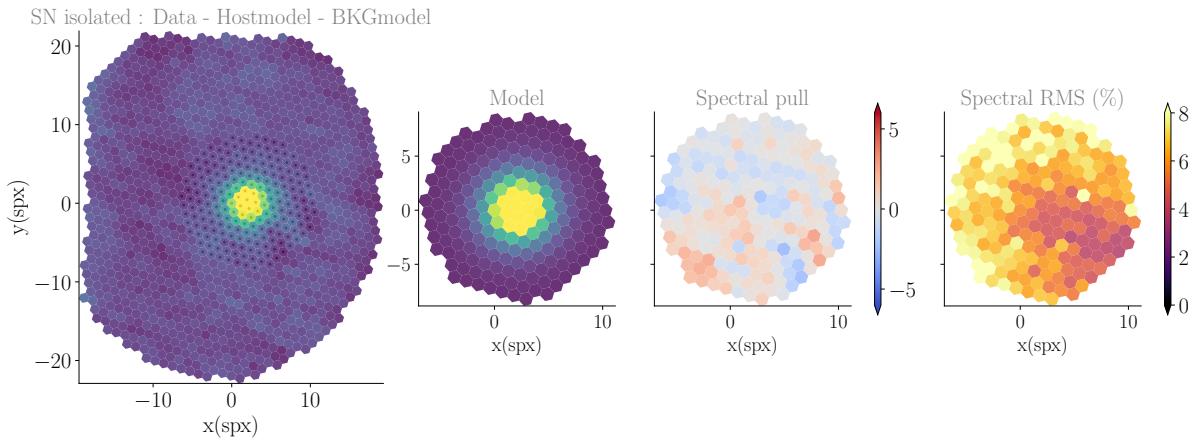


FIGURE 8.11 – Séparation de la supernova ZTF18accorf dans le cube SEDm. Chaque image de cube 3D correspond à une intégration entre 5000 et 8500Å. De gauche à droite : (a) le cube de données SEDm auquel nous avons soustrait le modèle de la galaxie et celui du fond, ce qui met en évidence la qualité d’isolation de la supernova et *a fortiori* la qualité de modélisation des deux autres composantes. (b) Le cube modèle de la source ponctuelle limité à une ouverture circulaire de 8 spaxels ($\sim 4''.5$) de rayon, définie par les spaxels mis en évidence dans (a) par les losanges noirs. (c) Le pull spectral et (d) le RMS spectral dans cette ouverture, tout deux définis comme dans la Figure 8.9. Nous pouvons voir ici l’absence de structure résiduelle, et un RMS spectral de l’ordre de 3-4% au niveau de la position de ZTF18accorf.

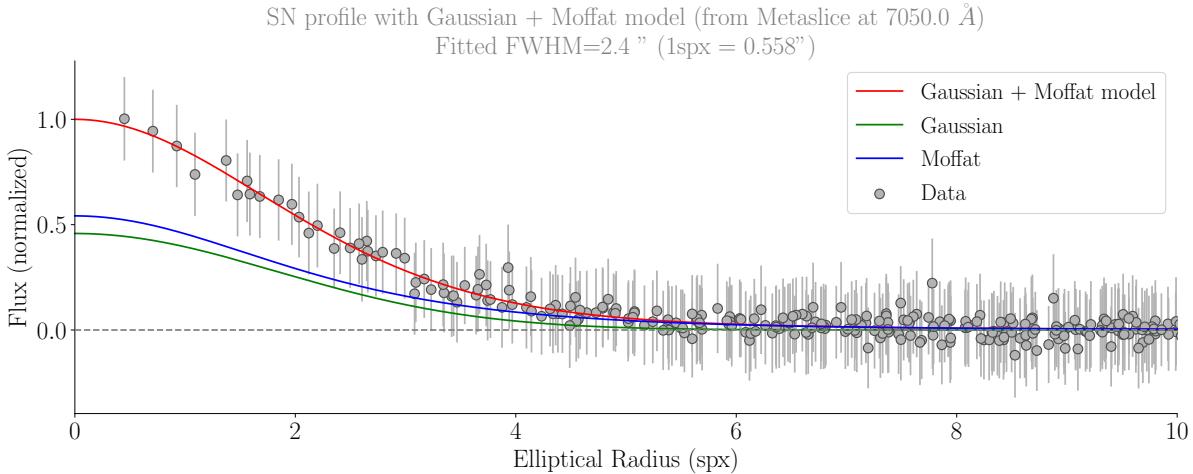


FIGURE 8.12 – Profil radial et modèle de PSF pour la méta-tranche à 7050Å de ZTF18accorf. Le pic du flux est ici normalisé à 1. Les points noirs indiquent les données de la méta-tranche du cube d’observation de la SEDm, après soustraction du modèle de la galaxie et du fond. Les courbes verte, bleue et rouge montrent respectivement l’ajustement de la composante gaussienne, Moffat et profil radial total du modèle de PSF pour cette méta-tranche. Le trait horizontal en pointillés indique un fond de ciel à 0 si cette composante a été parfaitement soustraite. Les ailes du profil tendant clairement vers cette valeur, cela nous conforte quant à la qualité d’ajustement de cette composante. Connaissant la taille d’un spaxel, nous pouvons également déterminer la largeur à mi-hauteur à cette longueur d’onde, ici 2''.4.

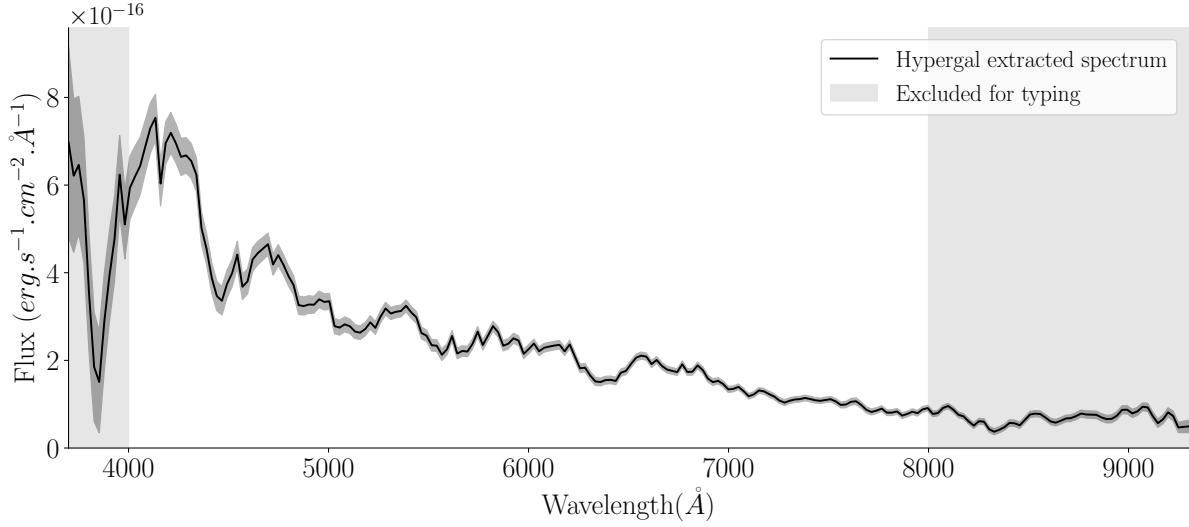


FIGURE 8.13 – Spectre extrait de ZTF18accorrf avec HYPERGAL. Tout le domaine spectral de la SEDm est affiché, et les bandes grises indiquent les zones qui ne sont pas utilisés pour la classification de la SN avec SNID.

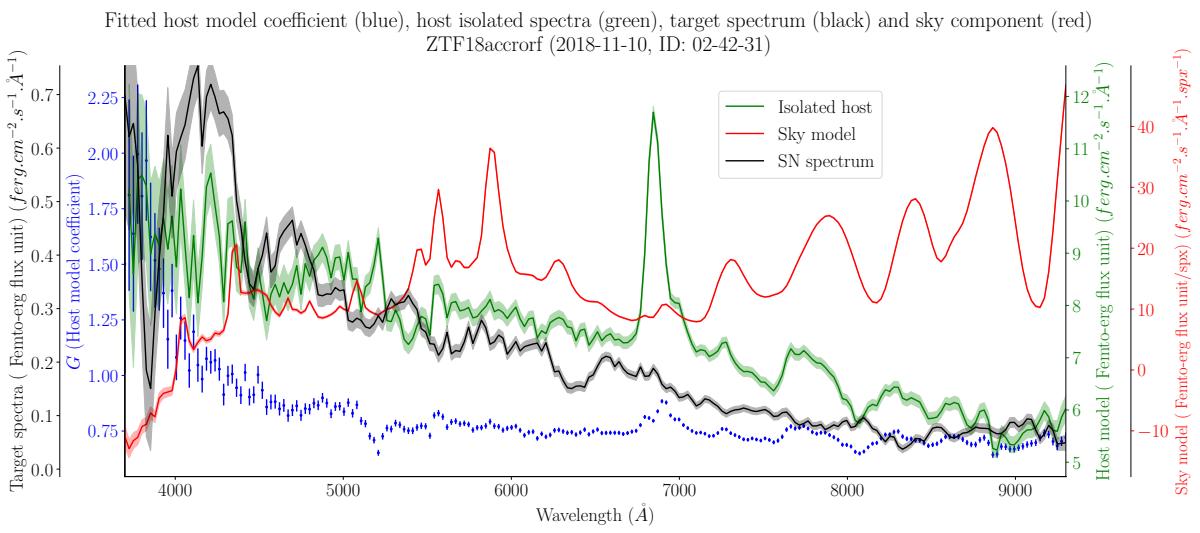


FIGURE 8.14 – Superposition du spectre des 3 composantes de la scène d'observation de ZTF18accorrf : fond de ciel en rouge, galaxie en vert, supernova en noir. Nous montrons également l'évolution chromatique de l'ajustement du coefficient de correction G du cube intrinsèque. Le but de cette superposition est de mettre en évidence l'absence de contamination évidente entre les spectres notamment par les raies d'émission de la galaxie. L'évolution du paramètre G montre l'adaptation partielle de l'intensité des raies O[III] et H α aux observations.

Par défaut, le pipeline est optimisé avec la librairie de calculs parallèles DASK¹ (?). Cet outil permet de sonder, avant de lancer les calculs, l’agencement des tâches et leur interdépendances, et de ce fait son utilisation optimise la parallélisation de code. Un autre avantage de DASK implémenté avec HYPERGAL est sa capacité à déployer le pipeline vers un cluster distribué en plusieurs *workers*. Nous avons configuré HYPERGAL pour être utilisable localement, mais également dans de tels cluster avec un simple argument. Les deux solutions d’ordonnancement de tâches disponibles sont SLURM (*Simple Linux Utility for Resource Management*) et SGE (*Sun Grid Engine*), permettant ainsi à HYPERGAL d’être déployé au Centre de Calcul de l’IN2P3 (CC-IN2P3). La Figure 8.16 illustre une partie de l’agencement des tâches du pipeline HYPERGAL, allant de la requête des données SEDm et PS1 jusqu’à l’ajustement des méta-tranches. Chaque bulle sur le schéma correspond à une tâche du pipeline (les sorties ne sont pas affichées), et les flèches illustrent les interdépendances. L’axe horizontal indique l’ordre de calcul, et l’axe vertical la parallélisation des tâches. Quelques étapes majeurs d’HYPERGAL sont indiquées par les boîtes grises. La parallélisation de certains calculs indépendants sont clairement illustrées, comme la requête simultanée des données PS1 et SEDm, la construction du cube intrinsèque et l’ajustement des conditions initiales sur les images photométriques, et enfin l’ajustement de toutes les méta-tranches. Les étapes suivantes (l’ajustement chromatique, l’ajustement linéaire des amplitudes qui parallélise les 220 tranches en longueur d’onde du cube de données, la reconstruction de la scène) ne sont pas montrées pour une question évidente de clareté. Toutes les sorties intermédiaires comme les figures de contrôle, les spectres extraits, les paramètres ajustés etc sont sauvegardées au format .hdf (*Hierarchical Data Format*). L’utilisation de DASK permet de réduire drastiquement le temps de calcul, passant de plusieurs heures sur un cœur (calcul linéaire) à ~ 20 min en utilisant seulement 10 coeurs. Plutôt que d’optimiser encore plus la rapidité de calcul pour une seule observation, nous avons concentrer nos effort sur l’optimisation du calcul simultané de nombreuses modélisations. La taille des clusters utilisables au CC-IN2P3 nous permet ainsi de faire tourner HYPERGAL sur une centaine de cubes en environ 1h. Lorsque la modélisation hyperspectrale n’est pas nécessaire, comme pour les étoiles standards, l’optimisation est encore plus impressionnante : les ~ 2000 étoiles standards dont nous avons extrait le spectre dans le chapitre ?? ont été traitées en seulement quelques minutes.

1. <https://www.dask.org>

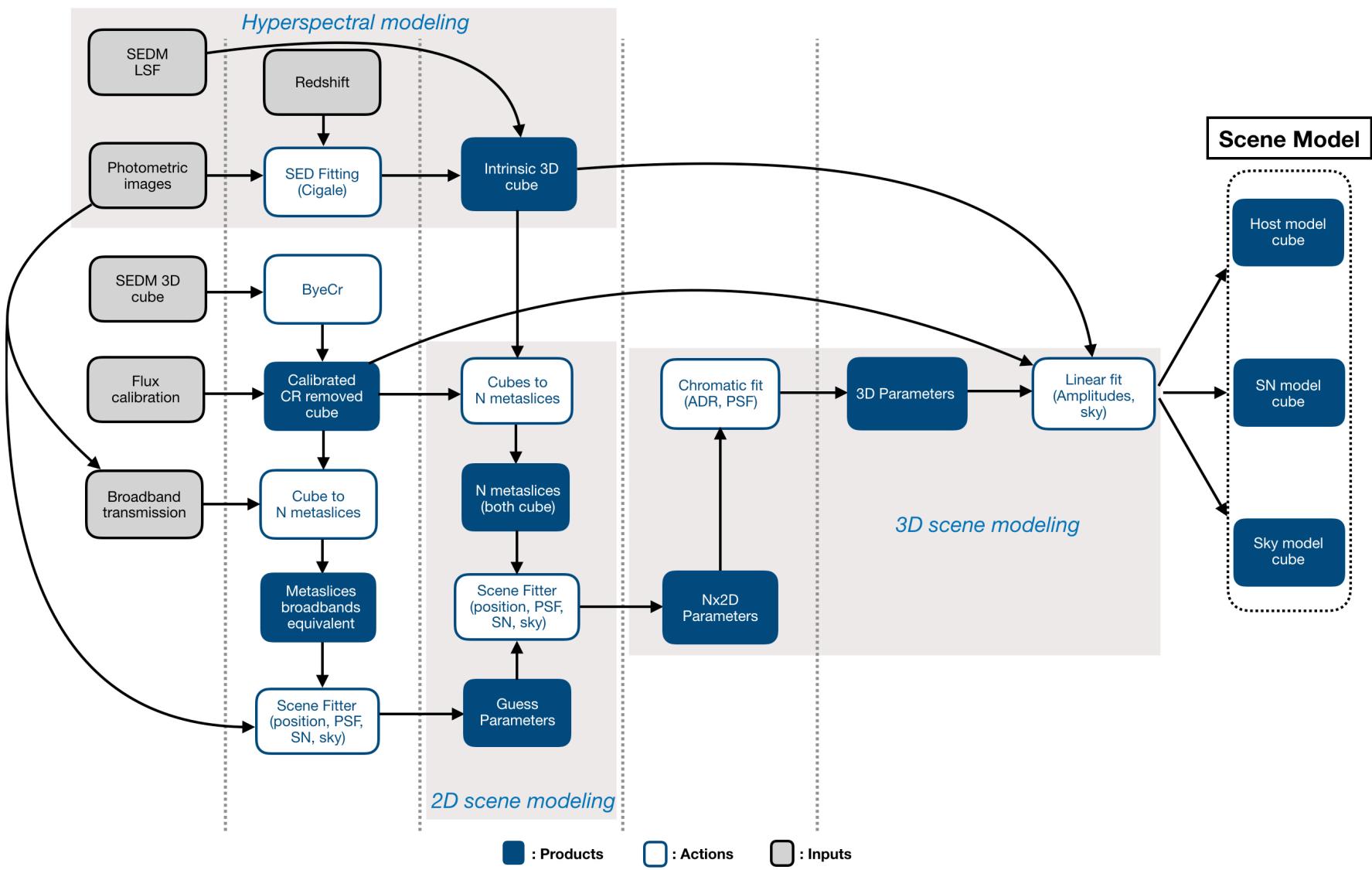


FIGURE 8.15 – Schéma complet du fonctionnement d’HYPERGAL. Les étapes sur la même verticale sont effectuées simultanément.

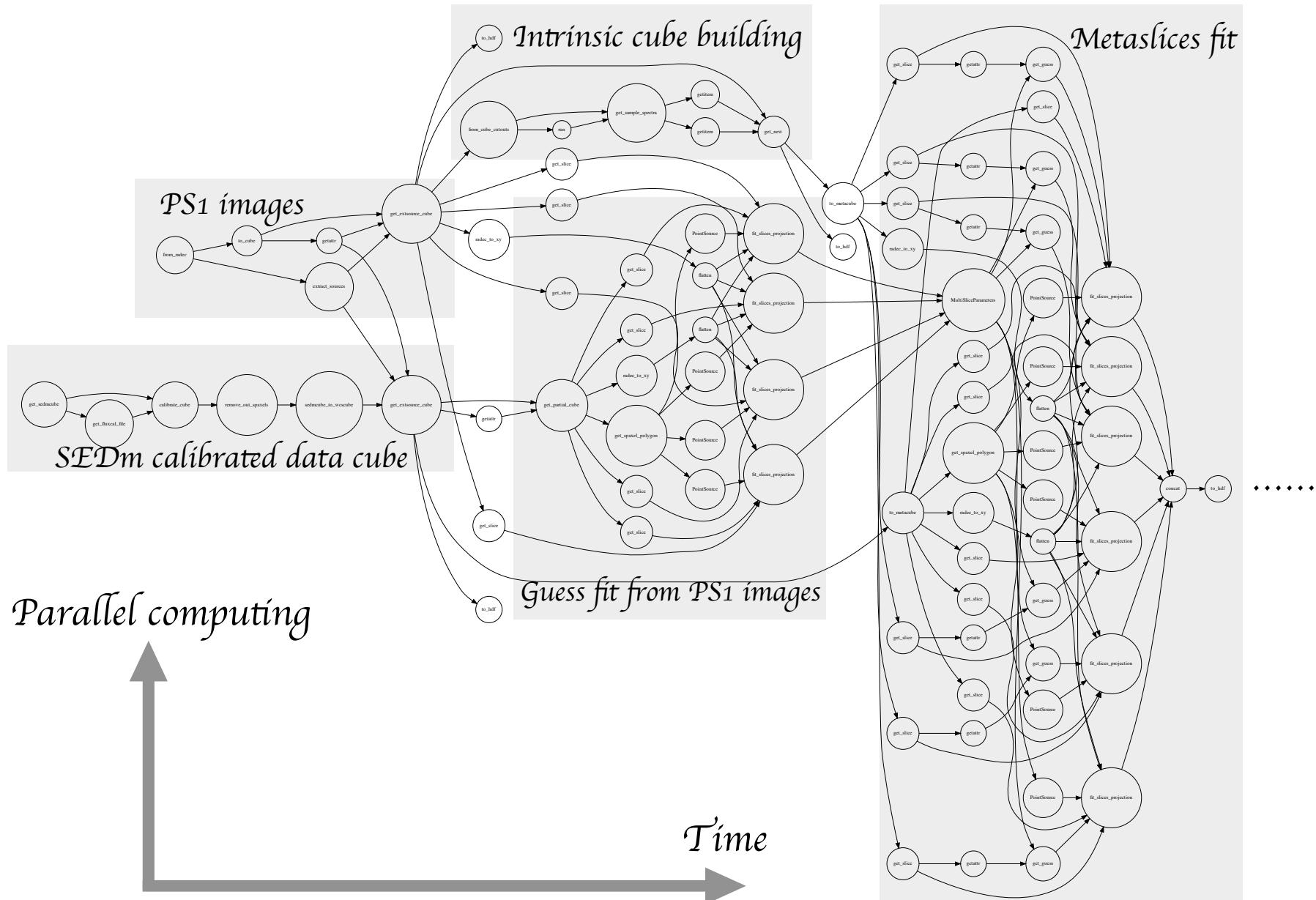


FIGURE 8.16 – Graphique d'une partie des calculs avec Dask et leurs interconnections du pipeline HYPERGAL, de la requête des données jusqu'à l'ajustement des N méta-tranches.

8.3 Classification : SNID

Le but principal du pipeline **HYPERRGAL** étant de permettre la classification spectrale des objets transitoires observés, nous avons intégré une méthode de classification automatique à l'image de celle utilisée initialement dans **PYSEDM**. Nous utilisons pour cela le software **SNID**, présenté dans le chapitre ???. Le domaine spectral utilisé par défaut pour la classification s'étend entre 4000 et 8000Å, les modèles utilisés n'étant généralement pas défini en dehors de ce domaine sachant la profondeur en redshift d'observation de ZTF ($z \leq 0.1$).

Afin de faciliter l'utilisation de **SNID** initialement écrit en **FORTRAN**, nous utilisons une *wrapper* **PYTHON** : **PYSNID**¹.

Nous montrons dans la Figure 8.17 la classification obtenue du spectre extrait de ZTF18accr0f avec **HYPERRGAL**, à comparer avec celle obtenue par extraction sans modélisation de la galaxie hôte par **PYSEDM**, que nous avons présenté dans la Figure ???. Initialement, la classification était incertaine, avec un paramètre de qualité *rlap* de 4.5 (le seuil minimal de "bonne qualité" étant 5), et un redshift estimé de $z = 0.178$ (au delà de la profondeur en magnitude de la SEDm). La classification avec le spectre extrait par **HYPERRGAL** montre pour le meilleur modèle un *rlap* de 14.7 pour une supernova de type Ia, à un redshift proche de celui utilisé pour la modélisation hyperspectrale de la galaxie hôte ($z_{snid} = 0.040$ contre $z_{in} = 0.042$). L'erreur sur le redshift pour un modèle avec **SNID** est de l'ordre de $\Delta z \approx 3.10^{-3}$.

Nous présentons également la distribution des 30 meilleurs modèles avec un *rlap* > 6, ainsi que le meilleur pour chaque sous-type de supernova pouvant correspondre au spectre extrait de ZTF18accr0f. Nous pouvons voir que seuls 2 sous-types semblent compatibles : une Ia-normal, ou une Ia-91T. Bien que la classification du sous-type soit plus subtile, la classification de ZTF18accr0f ne présente aucun doute comme étant une Ia. La phase de ce spectre est peu contrainte, les 30 meilleurs modèles étant ajustés avec une phase allant de -11 à +3 jours. Le redshift montre également une forte dispersion. Pour avoir une estimation approximative du redshift de la supernova à partir de la classification **SNID** des 30 meilleurs modèles, nous pouvons par exemple calculer le redshift moyen pondéré par le *rlap* (avec la condition que le *rlap* > 6). La déviation standard peut être estimer avec le nMAD. En faisant cela nous trouvons que $\bar{z}_{snid} = 0.045$ et un $nMAD(z_{snid}) = 0.005$.

8.4 Cas complexes

Bien que le cas utilisé pour présenter le pipeline **HYPERRGAL** (ZTF18accr0f) pouvait donner du fil à retordre à la méthode d'extraction directe (**PYSEDM**), nous considérons ce type d'observation comme idéal, avec une séparation assez nette de sa galaxie hôte.

Nous présentons dans cette section quelques cas d'extraction plus complexes de spectre avec **HYPERRGAL**, pour des observations où la supernova explose bien plus proche du centre galactique.

Pour toutes les observations que nous présentons ici, l'extraction avec le pipeline **PYSEDM** ne permet pas de classifier le spectre de la supernova, le cœur de la galaxie hôte étant extrait en même temps que la source ponctuelle (voir Figure ??).

¹. <https://github.com/MickaelRigault/pysnid>

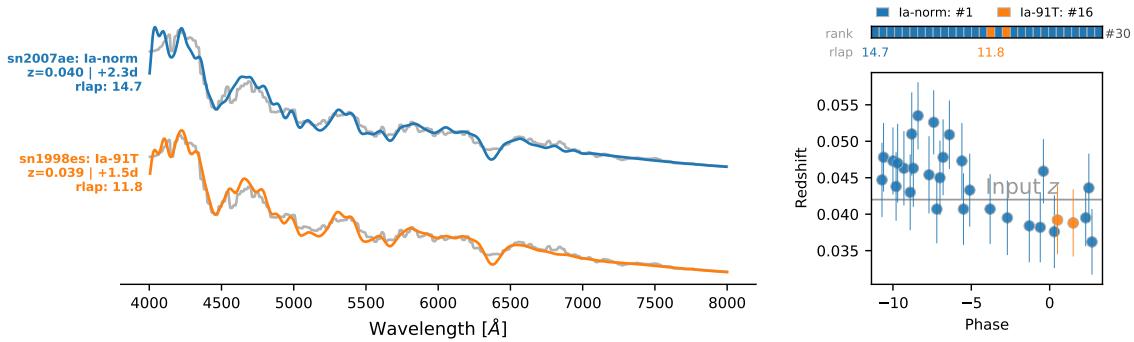


FIGURE 8.17 – Classification de ZTF18accorrf avec SNID. *À gauche* les modèles ayant le plus haut *rlap* pour chaque sous-type présent dans les 30 meilleurs modèles. Pour chacun nous y montrons le spectre extrait par HYPERGAL en gris, le modèle en couleur, le type et sous-type, le redshift, la phase et le *rlap*. *À droite* est présenté la distribution redshift/phase des 30 meilleurs modèles, avec pour le sous-type le même code couleur que pour les spectres affichés à gauche.

Nous présenterons successivement pour chaque observation : une visualisation globale de la modélisation de scène avec le RMS et le pull spectral à l’instar de la Figure 8.9 ; l’isolation de la galaxie hôte dans le cube de données SEDm et le spectre extrait dans l’ouverture définie dans la section 8.2.1 ; l’isolation de la supernova dans le cube de données SEDm et le spectre extrait (section 8.2.2) ; la classification avec SNID (section 8.3).

8.4.1 ZTF19acbjlnt

Cette supernova est celle que nous avons présenté à la fin du chapitre ??, afin justement d’illustrer le phénomène de contamination par la galaxie hôte lors d’un quasi-alignement avec le centre galactique dans la ligne de visée. La distance apparente entre cette supernova et le centre de la galaxie est d’environ $0''.8$, c’est à dire moins de la moitié de la largeur à mi-hauteur typique de la SEDm, et à un peu plus d’un spaxel de distance du centre du coeur de la galaxie hôte. Ce coeur étant étendu spatialement, la SN et la galaxie sont difficilement distinguables à l’oeil dans le MLA. La Figure 8.18 illustre une séparation assez remarquable de la SN et son hôte. Le RMS spectral de la modélisation de scène est de l’ordre de 5% dans la zone du MLA où les sources sont présentes. Cette supernova est classifiée par SNID comme une Ia avec un *rlap* de 17.6. La phase de ce spectre est peu contrainte mais semble indiquer une supernova proche de son pic de luminosité, allant pour les 30 meilleurs modèles de -8 à $+2$ jours. Le redshift d’entrée utilisé pour la modélisation hyperspectrale de la galaxie est $z_{in} = 0.059$, et en pondérant par le *rlap* nous trouvons un redshift moyen de $\bar{z}_{snid} = 0.060$ et $nMAD(z_{snid}) = 0.002$.

8.4.2 ZTF19abormno et ZTF20ablhllo

Les Figures 8.19 et 8.20 montrent deux autres extractions de supernovae (ZTF19abormno et ZTF20ablhllo) proches du centre de leur galaxie hôte, avec une séparation angulaire de $1''.5$ et $1''.2$ respectivement. Le RMS spectral de la scène à la localisation des sources dans le MLA est à nouveau d’environ 5%. La séparation des composantes de la scène est

également bien visible pour ces deux observations, et les deux objets sont classifiés par **SNID** comme des Ia avec un *rlap* pour le meilleur modèle de 21.9 et 13.2 respectivement. ZTF20ablhllo a la particularité d'avoir été observé très tardivement, présentant une phase de plusieurs dizaines de jours (entre +25 et +40 après le pic de luminosité).

Conclusion

L'association du cube intrinsèque de la galaxie hôte (chapitre ??) et de la réponse instrumentale de la SEDm (chapitre ??) a permis la conception d'**HYPERGAL**, un nouveau pipeline de modélisation de scène pour la SEDm.

HYPERGAL a été conçu pour répondre à la problématique de la contamination des supernovae par leur galaxie hôte, en ajustement simultanément toute les composantes de la scène observée. Avec cette méthode, chaque composante (fond de ciel, supernova et galaxie hôte) peut être extraite séparément, permettant son étude individuel. **HYPERGAL** effectue également une classification automatique du spectre extrait de l'événement transitoire étudié en utilisant le classifieur **SNID**.

Pour répondre également à un besoin d'efficacité et de simplification d'utilisation, cet outil est entièrement automatisé, optimisé par une méthode de parallélisation implémentée avec **DASK** et utilisable aussi bien localement que via un système d'ordonnancement de tâches comme *SLURM* ou *SGE*.

L'application de ce modéliseur de scène sur des supernovae très fortement contaminées par leur galaxie hôte montre un fort potentiel pour la classification d'objets inutilisables auparavant.

Afin d'avoir une idée des limites d'**HYPERGAL**, nous présentons dans le prochain chapitre une méthode de validation par le biais de simulations d'observations. Ce processus nous permettra de quantifier la précision spectrale d'extraction des SNe ainsi que l'efficacité des classifications, sachant *a priori* le type de l'objet simulé. Nous en profiterons pour appliquer l'ancienne méthode d'extraction (sans modélisation hyperspectrale de la galaxie hôte) sur ces modélisations, afin d'avoir une connaissance de l'amélioration qu'apporte **HYPERGAL**.

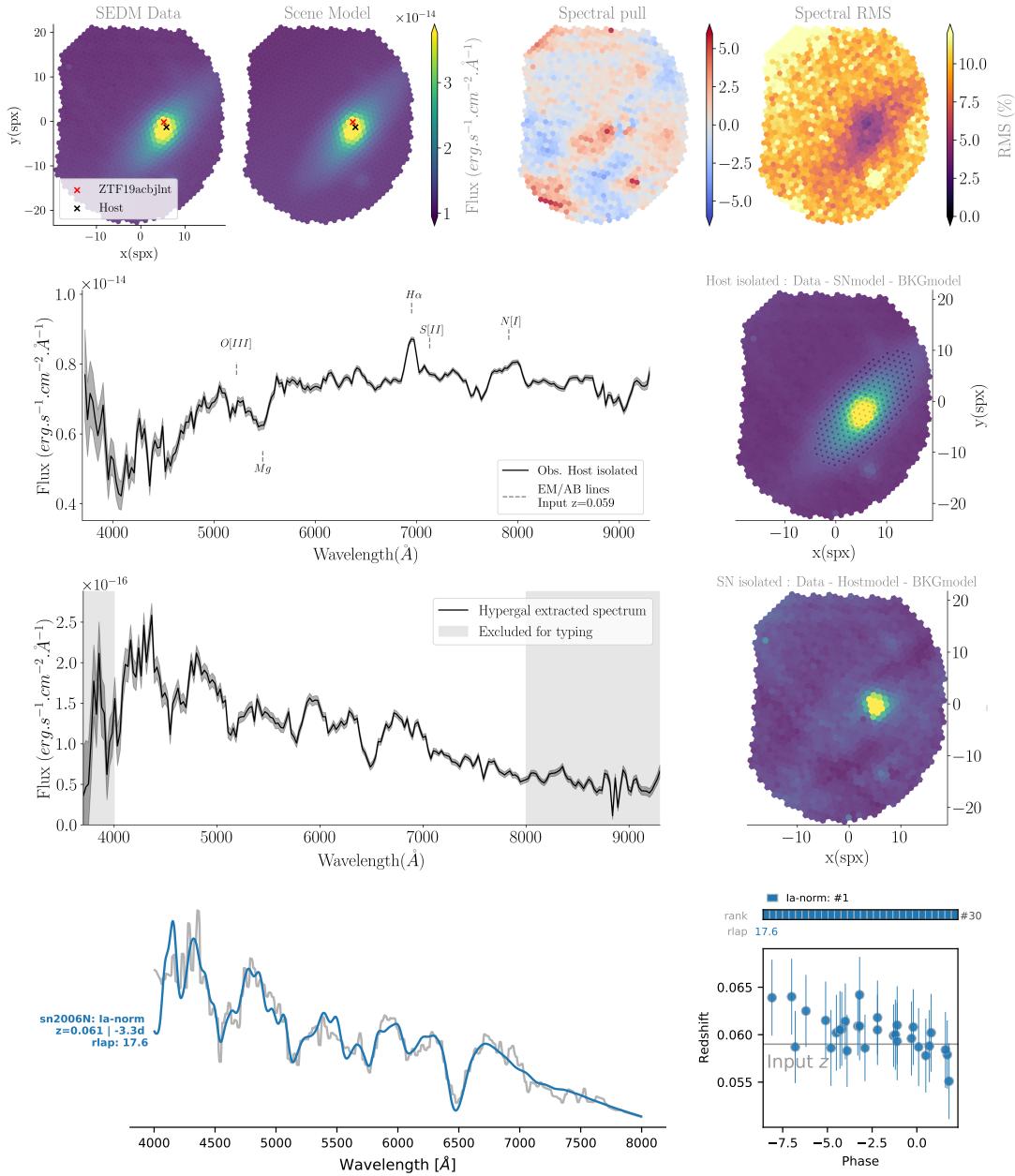


FIGURE 8.18 – Extraction de sources pour ZTF19acbjlnt avec HYPERGAL. *De haut en bas* : (a) visualisation globale de la modélisation de scène, pull et RMS spectraux, (b) l’isolation de la galaxie hôte, (c) l’isolation de ZTF19acbjlnt, (d) la classification SNID de ZTF19acbjlnt. La distance apparente entre la supernova et la galaxie est d’environ 0''.8 avec un redshift d’entrée pour la modélisation hyperspectrale de la galaxie de $z_{in} = 0.059$. Cette supernova est classifiée comme une Ia-norm avec un redshift moyen pondéré par le $rlap$ de $z_{snid} = 0.060$ et $nMAD(z_{snid}) = 0.002$.

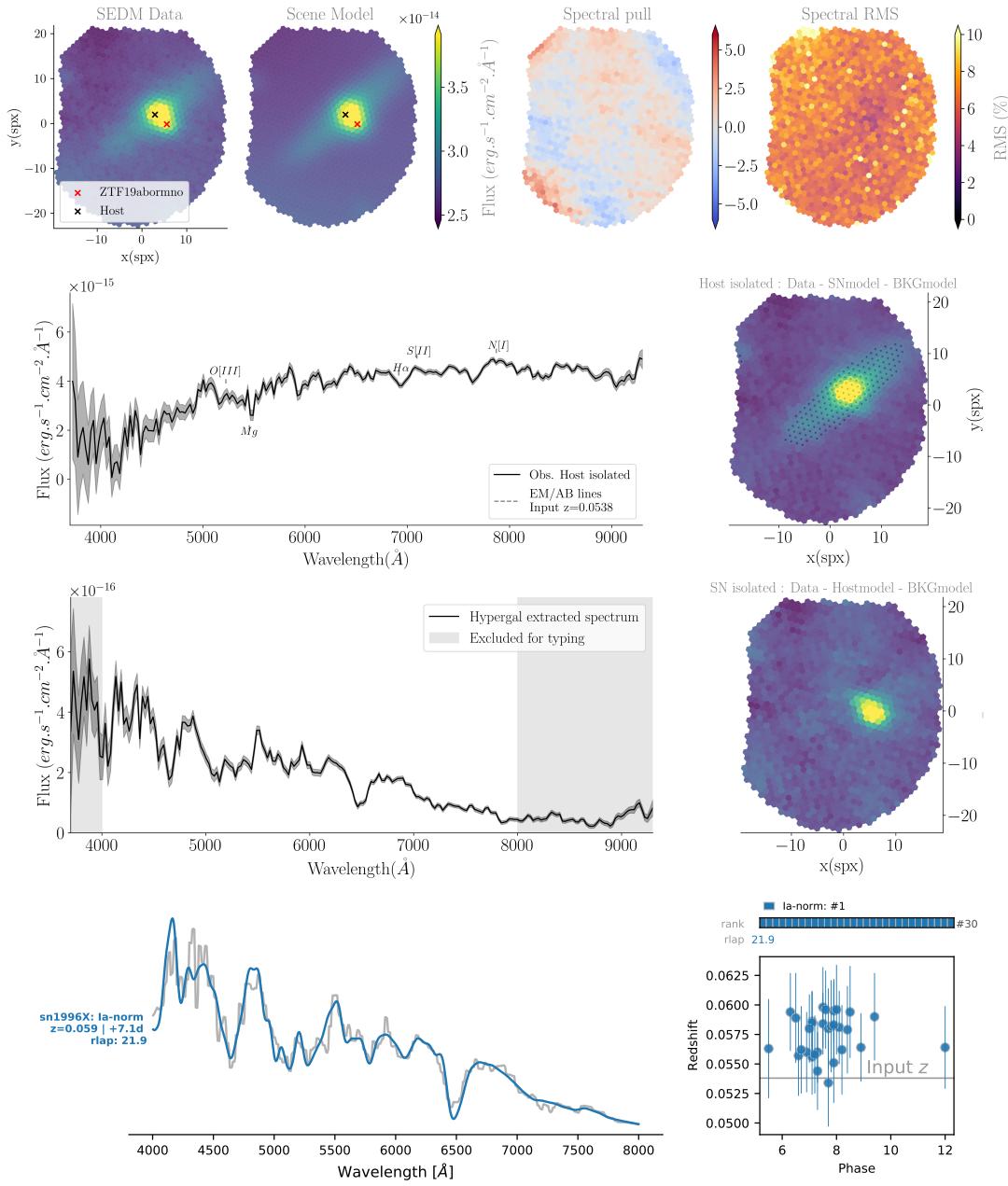


FIGURE 8.19 – Même chose que la Figure 8.18 pour la SN ZTF19abormno. La distance apparente entre la supernova et la galaxie est d'environ 1''.5 avec un redshift d'entrée de $z_{in} = 0.0538$. Cette supernova est classifiée comme une Ia-norm avec un redshift moyen pondéré par le $rlap$ de $\bar{z}_{snid} = 0.058$ et $nMAD(z_{snid}) = 0.002$.

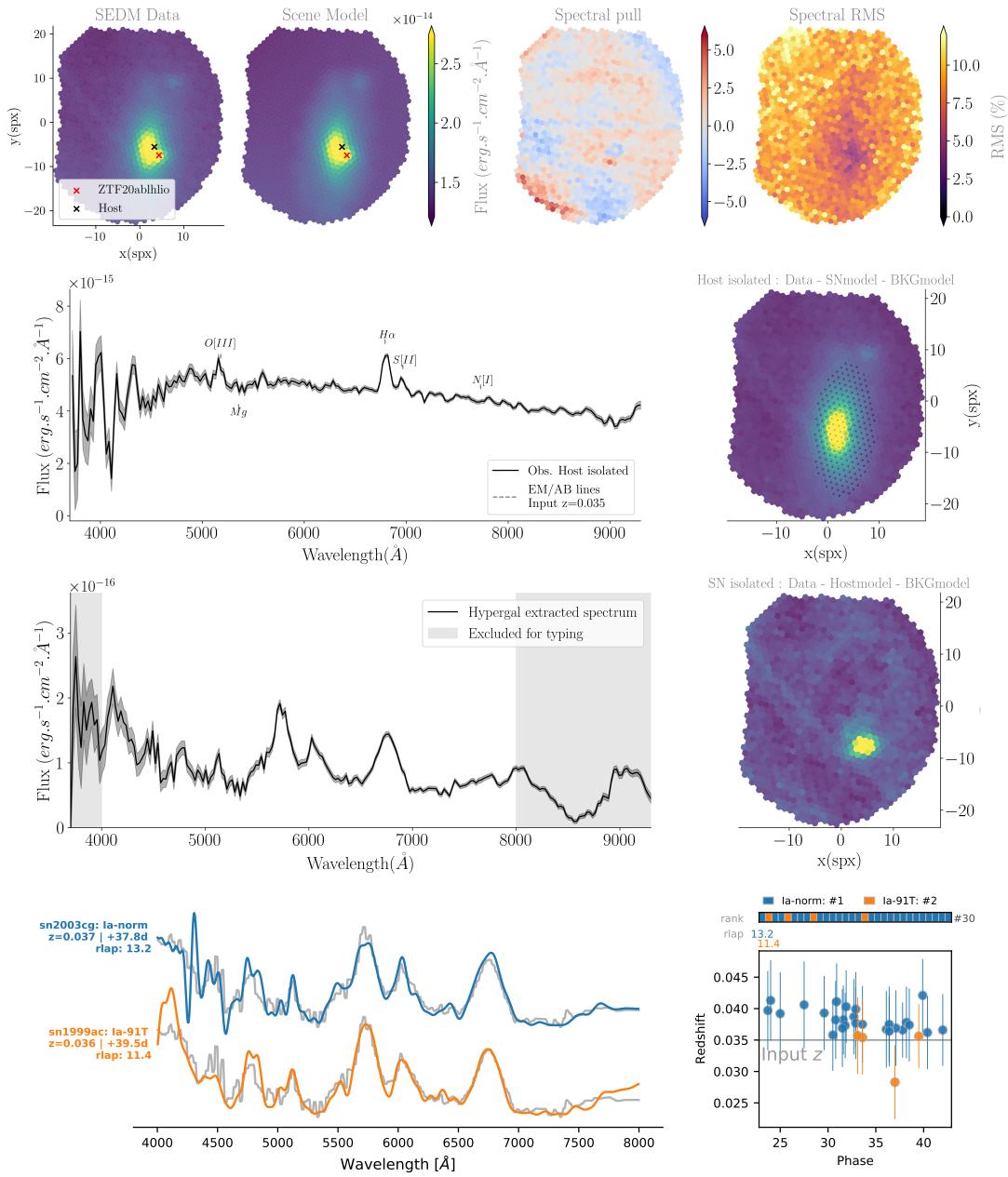


FIGURE 8.20 – Même chose que la Figure 8.18 pour la SN ZTF20abhlhlio. La distance apparente entre la supernova et la galaxie est d'environ $1''.2$ avec un redshift d'entrée de $z_{in} = 0.035$. Cette supernova est classifiée comme une Ia-norm avec un redshift moyen pondéré par le rlap de $\bar{z}_{snid} = 0.036$ et $nMAD(z_{snid}) = 0.001$.