# Data assimilation for aerothermal mean flow reconstruction using aero-optical observations: a synthetic investigation

Mohamed Yacine Ben Ali, Olivier Léon, David Donjat, Hervé Bézard, Emmanuel Laroche, Vincent Mons, Fréderic Champagnat

Association Aéronautique et Astronautique de France

*56th 3AF International Conference on Applied Aerodynamics*
*28 — 30 March 2022, Toulouse – France*

**FP74-AERO2022-benali**

# Data assimilation for aerothermal mean flow reconstruction using aero-optical observations: a synthetic investigation

**M.Y. Ben Ali[(1)], O. Léon[(1)], D. Donjat[(1)], H. Bézard[(1)], E. Laroche[(1)], V. Mons[(2)] and F. Champagnat[(3)]**
[(1)]*ONERA/DMPE, Université de Toulouse, France, mohamed_yacine.ben_ali@onera.fr*
[(2)]*ONERA/DAAA, Université Paris Saclay, France*
[(3)]*ONERA/DTIS, Université Paris Saclay, France*

## ABSTRACT

In this study, we investigate a Data Assimilation approach for the numerical reconstruction of turbulent aerothermal flows, and more particularly of their average density fields. The proposed approach relies on an ensemble Bayesian inference method providing a derivative-free optimization framework suitable for solving inverse problems and quantifying model uncertainties. The originality of this work lies in the use of non-intrusive and dense measurements such as those obtained by the optical technique referred to as Background Oriented Schlieren (BOS). Such a reconstruction problem is usually tackled with a purely aero-optical modeling as in tomographic BOS. The present approach introduces additional physical constraints by leveraging Reynolds Averaged Navier–Stokes flow models. To validate the relevance of this approach, the developed framework is applied as a simple calibration tool on numerical experiments involving a hot jet impinging on a flat plate, providing approximate inference for uncertain closure coefficients parametrizing a $k-\omega$ (SST) turbulence model.

## 1. INTRODUCTION

Aero-thermal flows characterized by in-homogeneous density fields are ubiquitous in aeronautical applications where heat exchanges can play an important role in the life-span of some mechanical components (cooled aero-engine blades for example) or in the aerodynamic performance of a structure (e.g. hot impinging jets for surface anti-icing). Having an accurate modeling capability of such aero-thermal flows would help manufacturers to develop optimal cooling or heating devices. Usually, experimental wind tunnel tests and numerical simulations are deployed separately in the investigation of such flows. On the one hand, numerical simulations may provide full fields of flow state quantities such as temperature, density or velocity. Among various techniques, Reynolds Averaged Navier–Stokes (RANS) simulations form a widespread and affordable approach. Yet, important limitations exist in RANS models that are usually detrimental for accurate simulations of aero-thermal flows. In particular, the level of heat exchange at a wall is generally unsatisfactorily predicted, which may require the use of higher-fidelity but more expensive simulations [6]. On the other hand, experimental techniques may provide reliable but limited information about such aero-thermal flows. For example, the Background Oriented Schlieren (BOS) technique [4], which exploits the refraction of light in the traversed flow to provide light deviation maps using a simple optical setup, can give some quantitative insights into the flow density field. This density field may then be reconstructed from the observed deviations based on a tomographic approach [13]. While such a technique can be efficiently used to study free flows, its application to bounded flows is more challenging because of the restrictions applying on the possible points of view. Furthermore, tomographic reconstructions are performed relying on an aero-optical modeling only and does not involve any prior or basic information on the flow of interest. Finally, it should be emphasized that this procedure can only provide estimates for the density field and not for other flow quantities such as the temperature or the velocity. Based on estimation theory, Data Assimilation (DA) [5] allows to overcome the respective limitations of numerical simulations and experiments, namely the lack of accuracy due to, among others, modeling errors for the former and the scarcity of flow information for the latter. In this study, we investigate an

ensemble-based Bayesian approach relying on filters developed for DA that supplies 2D maps of light-deviations to a RANS solver. The goal here is twofold. On the one hand, the present methodology aims at providing a full flow estimate from light-deviation measurements. On the other hand, it systemically enhances a baseline RANS prediction, which could not be used alone to accurately estimate the flow of interest.

The paper is organised as follows. Section 2 describes the methodological ingredients. Section 3, first, introduces the flow configuration, then discusses the results of the DA process with synthetic experimental data. Section 4 concludes the study and opens up with perspectives.

# 2. DATA ASSIMILATION FRAMEWORK

Given noisy BOS observations of an aero-thermal flow in the form of mean optical deviation maps [13], our objective is to infer an approximate mean flow state relying on a RANS model that naturally involves physical constraints on the flow quantities and that is considered in this first work to be subject to parameters uncertainty only. The present study thus aims at validating an approximate Bayesian model calibration method [10] that will be extended to account for model-form uncertainties [16] in future works.

## 2.1 A state augmented approach

The framework considered in the present study relies on a state-space model [1], and more particularly on a state augmentation approach recently proposed for inverse problems [8] and termed Ensemble Kalman Inversion (EKI) [3]. This approach relies on an iterated ensemble Kalman method [5] that blends observations and prior knowledge for estimation of the true flow state. In this sense, the present approach may be classified as a sequential DA technique were temporal dynamics is replaced by an artificial one, providing an approximate Bayesian inference method [16].

In this framework, the state-space model at iteration $K$ writes

$$X_K = \boldsymbol{M}(X_{K-1}, \eta^m), \tag{1a}$$

$$Y_K = \boldsymbol{H}X_K + \eta^o, \tag{1b}$$

with the $n$-dimensional state $X_K$ and the $m$-dimensional measurements $Y_K$. In a state-augmented framework, $X_K$ contains flow physical variables (and particularly $\overline{\rho}$ in our case), augmented with the uncertain RANS controling parameters. Eq. (1a) then models the artificial dynamics with $\boldsymbol{M}$ embedding the non-linear RANS operator that updates the state from iteration $K-1$ to iteration $K$ with some model uncertainty $\eta^m$. In the present numerical study, the model dynamics is considered as perfect and this system noise $\eta^m$ is neglected. Eq. (1b) represents the observation (or measurement) model in which $\boldsymbol{H}$ is a linear observation operator that projects the state $X_K$ to the measurement space with some measurement noise $\eta^o$. The linearity of the observation operator emerges from the type of measurement here considered and detailed in section 2.3. It is on this state-space model (1) that an ensemble-based Kalman Filter is applied to provide flow state and parameters estimates. Further details on the artificial dynamical model $\boldsymbol{M}$, on the observation model $\boldsymbol{H}$ and on the ensemble Kalman method used are provided in the following sections.

## 2.2 Aero-thermal RANS model

The aero-thermal RANS model constrains the estimated solution with conservation equations for mass density $\overline{\rho}$, momentum $\overline{\rho\mathbf{U}}$ and total energy $\overline{\rho E}$. The RANS models are closed by relations for the Reynolds stress tensor $\overline{\rho u_i' u_j'}$ for the dynamic part, and for the turbulent heat flux terms $\overline{\rho u_i' T'}$ for the thermal part. These two modeled terms entirely account for the unresolved physics and are the ones embedding uncertain parameters subject to optimization with the present approach.

Among the large variety of models available in the literature [15], we chose to use the $k-\omega$ SST (shear-stress transport) [11] turbulence model since it was reported in [6] to yield a very reasonable quantitative description of the flow considered in this work and detailed in section 3.1. Such a model is one of the eddy-viscosity models, which relies on Boussinesq's hypothesis. As mentioned previously, the associated model-form uncertainty is not accounted for in this work and we mainly consider model-parameters uncertainty for calibration. Such a calibration approach can be justified by the fact that model parameters are usually empirically determined by matching model results to some experimental or numerical results obtained for canonical flow configurations, and hence are not universal. We furthermore chose to only consider a few number of parameters and a sensitivity analysis with the flow configuration described in section 3.1 indicated the importance of at least four model parameters, namely $(\beta^*, \sigma_{k2}, \sigma_{\omega 2}, \beta_2)$. It can be noted that these chosen parameters are mainly acting on the outer part of the model (see [11]) that drives the free shear-flow region of the considered impinging jet flow as well as its turbulent/non-turbulent interfaces. Varying the inner part of the model was observed to induce convergence difficulties.

Closure for the RANS total energy equation is obtained in the present work using the simple gradient diffusion hypothesis to model the turbulent heat flux, such that

$$\overline{\rho u_i' T'} \simeq -\frac{\mu_t}{\mathrm{Pr}_t}\frac{\partial \overline{T}}{\partial x_i}, \tag{2}$$

where $\mu_t$ is the turbulent viscosity deduced from the $k-\omega$ SST model for the Reynolds stress tensor and $\mathrm{Pr}_t$ is the turbulent Prandtl number. This former parameter is the only one driving the thermal part of the present model and is thus added to the vector of uncertain parameters that finally writes $\Omega = (\beta^*, \sigma_{k2}, \sigma_{\omega 2}, \beta_2, \mathrm{Pr}_t)$. It may be noted that finer thermal models could have been considered, but the present basic modeling is here somewhat counter-balanced by the use of a Bayesian inference technique looking for optimal parameters.

Given a set of model parameters $\Omega_K$ at iteration $K$ in the framework defined in section 2.1 (with the other parameters of the RANS model left to their default values or deduced from explicit relationships), solving the RANS equations yields the $n$-dimensional flow state vector defined as

$$\chi_K = ((\rho_1, \mathbf{U}_1, P_1, k_1, \omega_1, T_1), \cdots, (\rho_M, \mathbf{U}_M, P_M, k_M, \omega_M, T_M)), \tag{3}$$

where $P$ denotes the mean pressure field and $M$ refers to the size of the computational grid. The over-bars and the subscript $\cdot_K$ are omitted on the flow state variables for clarity. As mentioned in section 2.1, this state is augmented to include the parameters $\Omega_K$, such that we define

$$X_K = (\chi_K, \Omega_K). \tag{4}$$

It is highlighted that the state solution given by this aerothermal RANS model with its "default" closure coefficients will be later-on referred to as the baseline state.

## 2.3 Observation model

The data $Y_K$ as given in the observation model (1b) results from the application of the observation operator $\boldsymbol{H}$ on the the flow state embedded in $X_K$. This operation involves performing a synthetic BOS observation [13] to evaluate 2D visualization maps of ray deviation angles in one projection plane (or camera). A multi-camera setup will be considered in future works dealing with non-axisymmetric flow configurations.

In a geometrical optics framework, the observed deviation angles result from the integral of the field of refractive index gradient $\nabla n_r$ along the path of light rays traversing the in-homogeneous fluid flow medium. The local refractive index is related to the fluid density by the Gladstone–Dale relation $n_r - 1 = G\rho$, where $G$ is a constant that depends on the gas composition (air in the present case) and the light wavelength. The three components of light ray deviations in space can then be written as

$$\varepsilon = \frac{G}{n_{r0}} \int_{s \in \mathtt{ray}} \nabla \rho \, \mathrm{d}s \tag{5}$$

where $n_{r0}$ is the refractive index of the medium surrounding the flow, which is assumed constant. As the ray path depends itself on the density, the relationship is generally nonlinear. A common way to simplify the problem is to use the paraxial approximation since the angles are usually small [13], providing a linear framework for the observation operator $\boldsymbol{H}$ that then amounts to a linearized ray-tracing operator. Furthermore, the three components $\varepsilon_{x,y,z}$ of the field of light ray deviations then approximately correspond to deviation angles and are thus considered as such in the following sections.

Given the preceding definitions, the observation data $Y_K$ in (1b) can be defined as

$$Y_K = ((\varepsilon_x, \varepsilon_y, \varepsilon_z)_1, \cdots, (\varepsilon_x, \varepsilon_y, \varepsilon_z)_L) \tag{6}$$

where $L$ is the number of light rays considered (or equivalently the number of points considered in the 2D deviation map). The dimension of $Y_K$ is $m = 3L$ and the observation operator $\boldsymbol{H}$ is of dimension $m \times n$ in matrix-form.

Finally, the experimental noise $\eta^o$ in (1b) is assumed to follow a Gaussian distribution $\mathcal{N}(0, \boldsymbol{R})$. The value of the (diagonal) co-variance matrix $\boldsymbol{R}$ was determined by a trial and error method to avoid convergence issues induced by possible negative model parameter values. This pragmatic approach is by no means optimal and proper ways of defining it and of thresholding parameters will be considered in future works.

## 2.4 Summary of the algorithm

Relying on the previously defined models and state-augmented framework, an approximate solution to the inverse problem is obtained following the steps given by Algorithm 1. Practical details on each step are provided in the following paragraphs.

**Algorithm 1** Iterative Bayesian inference procedure

**Initialisation**: $K \leftarrow 0$; generate initial ensemble $\{X_0^{(i)} \mid i \in \{1 \cdots N\}\}$;
**repeat**
(a) Calculate $\{X_{K+1}^{f,(i)}\}$ solving RANS equations,
(b) Perform observations $\boldsymbol{H} X_{K+1}^{f,(i)}$,
(c) Estimate $\{X_{K+1}^{a,(i)}\}$ using Ensemble-based KF, update $K \leftarrow K+1$
**until** Exit convergence criterion.

**Initialization** A prior distribution of the model state is required for initialization of the ensemble-based procedure. In this study, different realizations of the state are simply obtained by perturbing the set of model parameters gathered in $\Omega$ from the baseline. A Latin hyper-cube sampling method is selected to draw near-random values for the parameters in a range of $\pm 50\%$ of their baseline values, providing an ensemble of initial parameters $\{\Omega_0^{(i)} \mid i \in \{1, \cdots, N\}\}$ with $N$ the number of ensemble members set to 50 for the considered experiments. The

associated ensemble of initial flow states $\{\chi_0^{(i)}\}$ is then evaluated using the aero-thermal RANS model described in section 2.2 to build the ensemble of initial augmented states $\{X_0^{(i)}\}$.

**(a)** RANS simulations are performed for every member of the ensemble of parameter values estimated in the preceding inference iteration. These simulations are carried out using the ONERA multi-physics platform CEDRE [14]. The compressible RANS equations are discretized by a finite volume method on a unstructured mesh. Convergence of each RANS simulation are ensured with a stopping criterion on the residual of conservative quantities that are required to drop by at least 1 order of magnitude.

**(b)** The 2D BOS synthetic observations are deviations fields evaluated from numerically-obtained 3D density flow fields (obtained from RANS simulations or LES results in the current study, see section 3). These optical deviations are obtained by considering a synthetic camera whose spatial location with respect to the flow field and whose optical properties are first specified. A dedicated ray-tracing algorithm developed for tomographic BOS [13] and coded to work on GPUs are used.

**(c)** An Ensemble Transform Kalman filter (ETKF) [2] is used in the estimation of optimal values of turbulence model parameters. Among others, this method is more suitable for nonlinear models and high dimensional configurations than classical Kalman filter [9]. The open-source project PDAF (Parallel Data Assimilation Framework [12]) is used to perform the ETKF analysis step.

**Exit criteria** This ensemble-based inference methodology is iterated to obtain converged estimates and no stopping criterion are used in this validation work. Yet, for practical applications with time and resources constraints, various criteria may be considered. For instance, one may use the $L_2$ norm of the relative variations of the parameters, the same norm for the observed quantity, or also the relative reduction on Mean Root Mean Square Error (MRMSE) on the state that will be discussed in the following results.

# 3. RECONSTRUCTION OF SYNTHETIC EXPERIMENTS

This section presents calibration results obtained with synthetically generated data. These synthetic data provide known ground-truth to assess the validity of the present inference method. The application case investigated is a subsonic hot-jet flow impinging a flat plate.

## 3.1 Flow configuration

This impinging jet flow configuration is illustrated in Fig. 1. A fully-developed turbulent air flow heated at a total temperature $T_J$ of 130 °C is issued from a round pipe of diameter $D = 0.06\,\mathrm{m}$ at $x/D = 0$ with a centerline mean axial velocity $U_J = 31.5\,\mathrm{m\,s^{-1}}$. The bulk Reynolds number is $\mathrm{Re}_b \approx 60,000$. The subsonic turbulent hot jet impinges a flat plate located at $x/D = 3$ and placed perpendicular to the jet axis. The simulation domain radially extends up to $y/D = 12$ to capture the development of the induced wall jet. This turbulent flow is statistically symmetric around the jet axis and 2D axisymmetric RANS simulations were performed. Relying on a previous extensive experimental study [7], boundary conditions for the jet mean axial velocity profile in the pipe exit plane and for the entrained air flow mean velocity profile at $x/D = -3$ are set using LDV measurements. Furthermore, the plate temperature distribution $T_P$ is defined using infrared temperature measurements obtained at thermal equilibrium. The simulations are performed using an unstructured mesh containing approximately $42,000$ grid cells. Time integration is carried out using a first-order implicit scheme and a first-order spatial scheme is used to ensure stationarity of the solutions for the range of parameters explored. Synthetic BOS observations are performed with one camera oriented with its optical axis parallel to the $z$ axis. The observation resolution is chosen to provide dense deviation maps with a dimension of $(n_x = 500, n_y = 500)$, thus such that the size of observation data $Y_K$ is $m = 3 \times 2.5 \times 10^5 = 7.5 \times 10^5$.

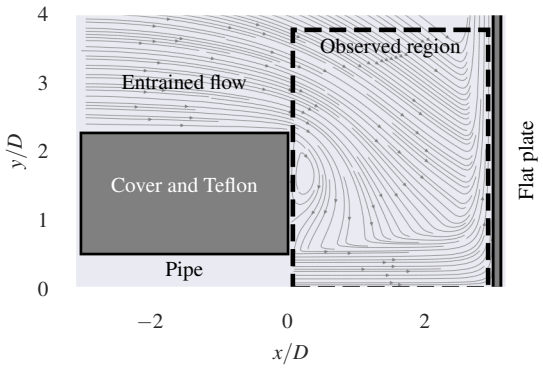



Figure 1: Computation domain for the hot-jet flow impinging on a flat plate.

## 3.2 Numerical twin experiments

The validity of the present framework for the estimation of RANS model parameters using BOS observations was first evaluated using synthetic data from two numerical twin experiments performed with the RANS model itself. This is an idealized setup where the model used

Table 1: Mean values and standard deviations before ("prior") and after ("estimate") estimation by the iterative Bayesian inference method along with the pseudo-true values of the three parameters chosen for the two twin experiments.

(a) Twin experiment $\text{RANS}_1$.

| | $\beta^*$ | | $\sigma_{\omega 2}$ | | $\beta_2$ | |
|---|---|---|---|---|---|---|
| Statistic | Mean | Std. | Mean | Std. | Mean | Std. |
| Prior | 0.09 | $2.6e-02$ | 0.856 | $2.5e-01$ | 0.0828 | $2.4e-02$ |
| Estimate | 0.1076 | $2.7e-05$ | 0.658 | $3e-04$ | 0.1029 | $3.4e-5$ |
| Pseudo-true | 0.1076 | - | 0.663 | - | 0.103 | - |

(b) Twin experiment $\text{RANS}_2$.

| | $\beta^*$ | | $\sigma_{\omega 2}$ | | $\beta_2$ | |
|---|---|---|---|---|---|---|
| Statistic | Mean | Std. | Mean | Std. | Mean | Std. |
| Prior | 0.09 | $2.6e-02$ | 0.856 | $2.5e-01$ | 0.0828 | $2.4e-02$ |
| Estimate | 0.0455 | $1.1e-04$ | 0.694 | $1.4e-02$ | 0.0887 | $5e-04$ |
| Pseudo-true | 0.0459 | - | 0.772 | - | 0.0942 | - |

can perfectly match the experimental observations, thus for which no model-form uncertainty exists. In these numerical experiments, some values of the RANS model parameters were arbitrarily chosen and were used to generate synthetic data considered as experimental observations. These values are later-on referred to as the pseudo-true values. It is emphasized that these cases of study do not provide realistic solutions and should only be considered as artificial experiments. For this twin experiments study, only the three parameters $(\beta^*, \sigma_{\omega 2}, \beta_2)$ are considered to simplify the analysis and to restrict the problem to a purely aerodynamic one.

In the first twin experiment, named $\text{RANS}_1$, the parameters were chosen to be close to their baseline values, within a $\pm 25\%$ range. This provides a first simple test-case with pseudo-true values within the prior distribution. Baseline and pseudo-true values are given in Table 1a. The second experiment, named $\text{RANS}_2$, explores the case were the state to recover is more different, with parameters further away from the baseline as shown in Table 1b. Notably, a maximum relative difference of about 50% is set on $\beta^*$, that is almost out of the prior distribution. The pseudo-true density fields obtained for the twin-experiments $\text{RANS}_1$ and $\text{RANS}_2$ are given in Fig. 2(a) and Fig. 3(a) respectively, illustrating the qualitative difference between the two cases, in comparison with the initial baseline state (Fig. 2(b) and reproduced in Fig. 3(b)). Finally, the initial data-model discrepancies in terms of root mean square error (RMSE) between the pseudo-true state and the initial baseline state for the two twin-experiments are displayed in Fig. 2(c) and Fig. 3(c). This (normalized) RMSE between two states, $\phi$, is defined as

$$\text{RMSE}(\phi_{\text{ref}}, \phi_k) = \frac{\sqrt{(\phi_{\text{ref}} - \phi_k)^2}}{\max(\phi_{\text{ref}}) - \min(\phi_{\text{ref}})}. \qquad (7)$$

As shown in these figures, the initial discrepancies reach maximum values of about 10% of the maximum density variations in the pseudo-true state for $\text{RANS}_1$ and up to 40% for $\text{RANS}_2$.

These discrepancy maps provide indications on the differences of density fields between the baseline (prior sample mean) and the pseudo-true states, but state observations are actually performed on BOS deviation fields. It is then instructive to examine the associated discrepancies in measured light deviations, as shown in Fig. 4 for $\text{RANS}_1$. Only the RMSE for two components of the deviation field are shown ($\varepsilon_x$ and $\varepsilon_z$) since the third one ($\varepsilon_y$) is one order of magnitude smaller. First of all, these maps emphasize that dense observations are obtained on the entire jet flow. Second, since $\varepsilon_x$ provides information about differences in density gradients along the $x$ axis, the main errors are naturally concentrated in the impingement region and in the wall jet. Discrepancies on $\varepsilon_z$ however mainly reflect differences in the jet shear-layers. Finally, the normalized error amplitudes are of the same order of magnitude (here about 8% on $\varepsilon_x$) as the relative discrepancies on the density field in Fig. 2(c).

Calibrations were then performed on these two test-cases. Fig. 5 shows the convergence of the spatial average of RMSE, denoted MRMSE, in log-scale on the observed deviations for the two experiments. It is highlighted that these global relative discrepancies on the sample means significantly decrease to less than 0.1% of the baseline maximum deviation difference within 4 iterations.

Convergence of the samples and of the sample-mean values of the model parameters is furthermore shown in Fig. 6. The final estimates are given in Tables 1a and 1b that list the means and standard deviations of the parameters before and after inference, together with the corresponding pseudo-true values. From these results, it can be emphasized that mean estimates perfectly match the pseudo-true values of the experiment $\text{RANS}_1$ (less than 1% error on all the parameters) while a satisfactory match is obtained for the more distant case $\text{RANS}_2$. Indeed, a maximum 11% error on $\sigma_{\omega 2}$ is obtained, but interestingly this parameter is also identified as less sensitive as the estimate is associated with a larger sample standard-deviation (Std.) in Table 1b. The observed convergence on the parameters estimates is associated with span reduction of the ensemble members in Fig. 6, with standard deviations reduced by 1 to 3 orders of magnitude depending on the parameter. Regardless the experiment, $\beta^*$, the coefficient involved in turbulence dissipation, was retrieved with at least 2 significant digits as shown in Table 1a. Compared to the other two considered parameters, reduced standard deviations are obtained, thus pointing out its major role in the dynamic part of the present RANS model.

It may be highlighted here that running the iterative inference method on the experiment $\text{RANS}_2$ required an

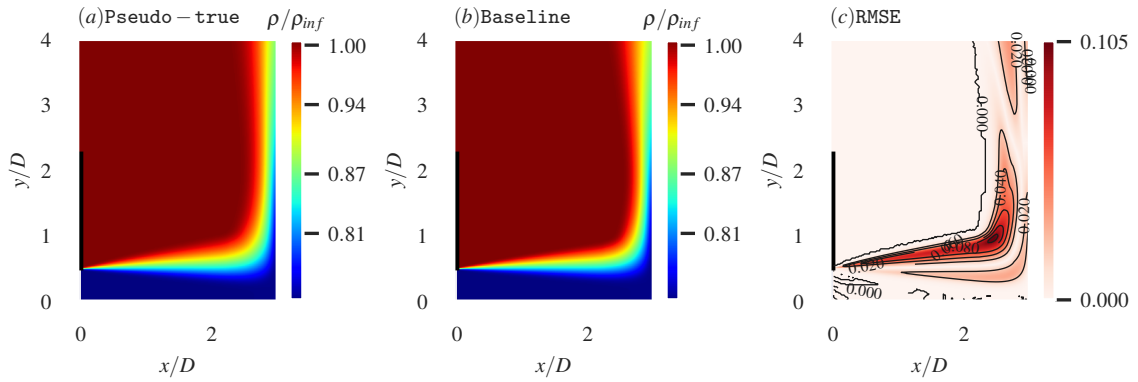



Figure 2: Density distributions of (a) the pseudo-true state for the twin experiment $\text{RANS}_1$ and of (b) the baseline solution (also corresponding to the prior samples mean). (c) Map of $\text{RMSE}(\rho_{\text{p-t}}, \rho_{\text{bsl}})$ between the two previous fields.

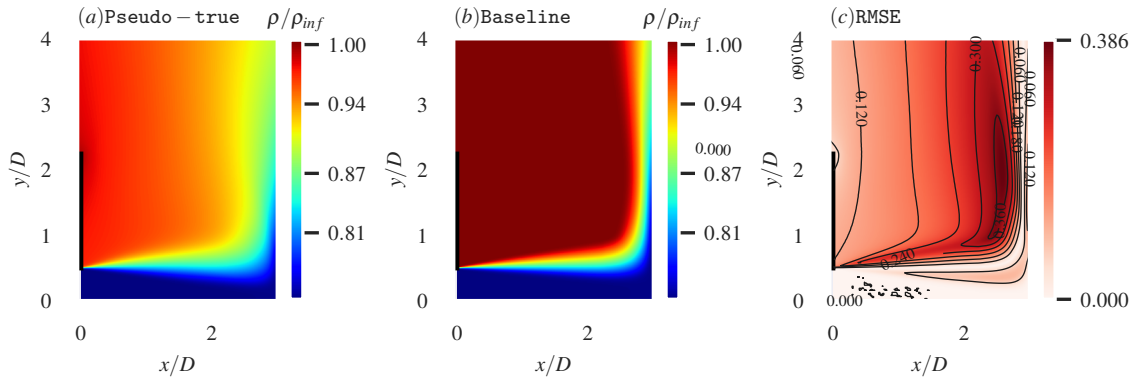



Figure 3: Density distributions of (a) the pseudo-true state for the twin experiment $\text{RANS}_2$ and of (b) the baseline solution (also corresponding to the prior samples mean). (c) Map of $\text{RMSE}(\rho_{\text{p-t}}, \rho_{\text{bsl}})$ between the two previous fields.

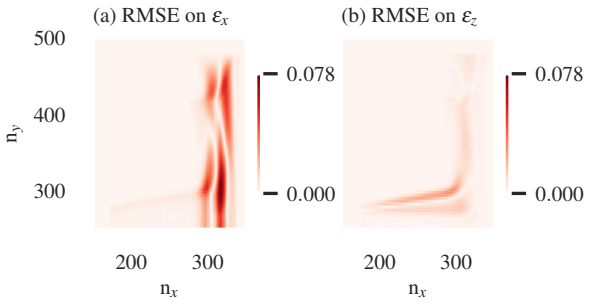



Figure 4: 2D maps of RMSE of observed deviations between the pseudo-true state and the baseline state for $\text{RANS}_1$.

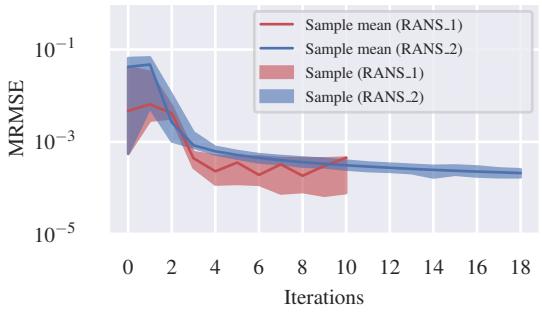



Figure 5: Convergence of $\text{MRMSE}(Y_{\text{p-t}}, Y_K)$ during the assimilation procedure for the two twin experiments. $Y_{\text{p-t}}$ refers to the pseudo-true deviations.

empirical reduction in the variance of $\boldsymbol{R}$, as already emphasized in section 2.3. Such an approach yielded realistic estimates but it also brought less confidence in the experimental data and resulted in a apparent slow convergence on the ensemble members. It nevertheless provided more stability to the process and a narrower spread of the MRMSE on BOS measurements as shown in Fig. 5.

Given the relative accuracy of the estimated model parameters, the resulting density flow field for the experiment $\text{RANS}_2$ is now examined, this test-case being less favorable than $\text{RANS}_1$. Comparing Fig 3 and Fig. 7, the resulting mean estimate of density field has been drastically improved, with maximum RMSE values of about 0.5%.

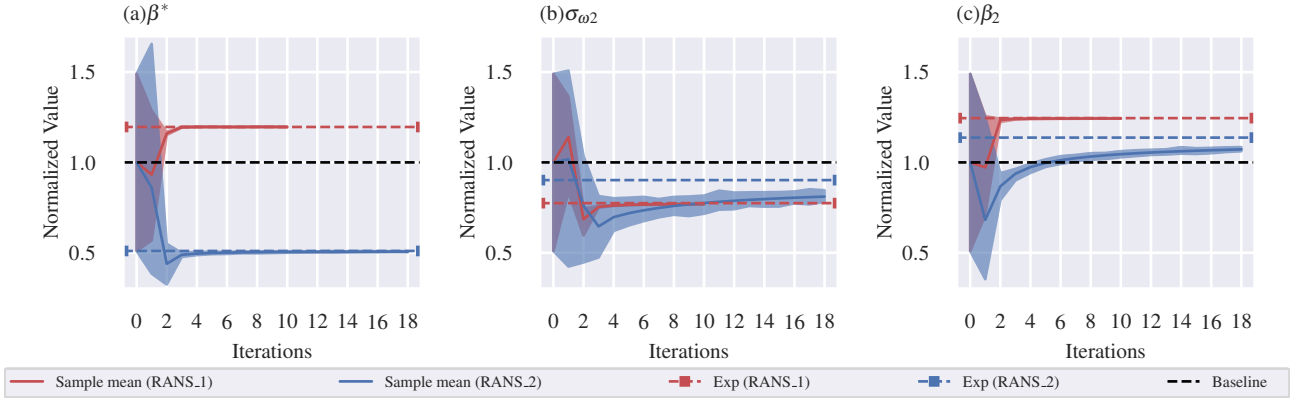



Figure 6: Convergence of the three model parameters during the assimilation procedure for the two twin experiments, for the sample and for the sample-mean. The values of the parameters are normalized by the baseline values.

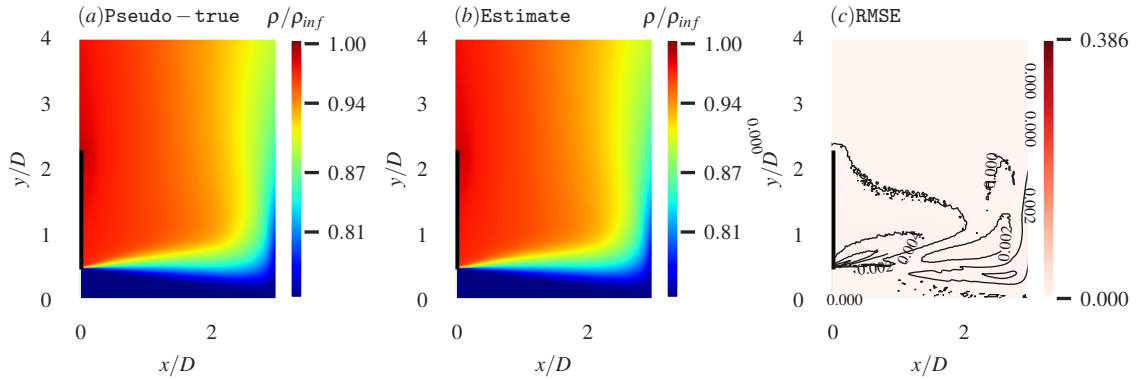



Figure 7: Density distributions of (a) the pseudo-true state for the twin experiment $\text{RANS}_2$ and of (b) the posterior solution obtained by the ensemble-based method (sample-mean at iteration 18). (c) Map of $\text{RMSE}(\rho_{\text{p-t}}, \rho_{\text{K=18}})$ between the two previous fields.

These first results provide support for the relevance of the present ensemble-based inference method of RANS models and approximate reconstruction of density fields from 2D maps of BOS measurements only. However, as previously emphasized, the idealized synthetic data here considered were obtained relying on the RANS model used for inference and challenges induced by model-form inadequacy should be expected with more realistic flow measurements.

## 3.3 Application to synthetic LES data

To investigate the case where more realistic flow observations are available, the present calibration approach was applied on synthetic data obtained by Large Eddy Simulation (LES) on the same impinging jet configuration [6].

In this study, two sets of model parameters are considered: a first set of 3 "aerodynamic" parameters as in the previous twin experiments and referred to as "aero only" in the following figures; and the full set of 5 model parameters $\Omega$ defined in section 2.2 to better account for thermal effects, referred to as the "full set" later-on.

Applying the present calibration method yielded the convergence history on the MRMSE of the BOS observations presented in Fig. 8. For both sets of parameters, the relative global errors decreased below 1% and reached a plateau after $7-8$ iterations. While no significant differences appear on this global result (where a slightly better reduction was achieved with the full-set), the estimated values of the ensemble members actually converge to different values for the two considered sets of parameters. This is made evident in Fig. 9A and Fig. 9B. Particularly, the inclusion of the parameters related to the RANS thermal model significantly modified the convergence values of the parameters $\sigma_{\omega 2}$ and $\beta_2$. Interestingly, regardless of the chosen set of parameters, the ensemble values of $\beta^*$ did not diverge much from the baseline. Furthermore, examining the ensemble spread, it can be observed that it also led to some increase of the variance of $\sigma_{\omega 2}$ and $\beta_2$, especially for the later parameter. This suggests that the parameters related to the thermal model might offer a higher degree of freedom to the estimated state. A

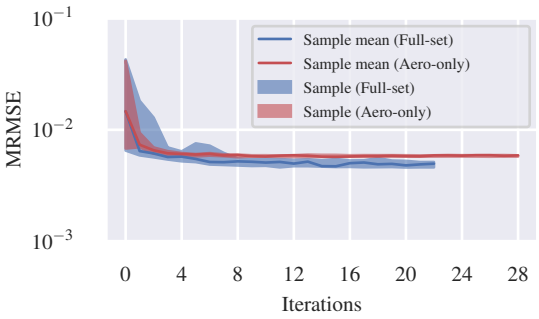



Figure 8: Convergence of MRMSE$(Y_{\text{p-t}}, Y_K)$ during the assimilation procedure of the LES data for the two sets of chosen parameters.

finer investigation of the uncertainty within the thermal-model-part is yet to be addressed in a future work. Regarding the two parameters $\text{Pr}_t$ and $\sigma_{k2}$, reductions from their initial baseline values are observed, suggesting an enhancement of temperature diffusion and turbulent kinetic energy diffusion compared to the baseline.

The previously observed convergence on the model parameters and on the global observation error can also be confirmed by comparing prior and posterior density fields, displayed in Fig. 10 and Fig. 11 respectively, along with the true-state (LES field) and the associated RMSE maps. Significant improvements in terms of relative errors are obtained. With the prior baseline state, maximum relative errors of about 35% were found in the jet shear layer in Fig. 10(c), while maximum errors between $5-10\%$ are obtained with the posterior solution in Fig. 11(c). More particularly, the wall jet is almost perfectly reproduced and most of the discrepancies are found in the initial jet shear layer.

Density profiles for prior and estimate samples are furthermore presented in Fig. 12 and Fig. 13, along with the LES results. As observed in Fig. 12, the baseline and the prior samples clearly deviate from the LES results in terms of jet shear layer thickness. Despite showing a larger radial diffusion, the estimated density profiles also significantly deviate from the true state in this region. In the wall jet region displayed in Fig. 13, however, the density profiles of the estimate sample mean are in much better agreement compared to the baseline results. The remaining local discrepancies observed between the estimate samples mean and the LES data are likely the result of model inadequacies, the proposed aero-thermal RANS model being very rigid. It can be concluded that the LES mean density field does not reside in the space spanned by the prior ensemble since the present approach can only find approximate solutions within this span [8]. Among the limitations of the present approach, we particularly emphasize that only global model parameters were considered in the assimilation process while the studied flow shows at least two different regions (free jet and wall jet), where local model adaptations may be more relevant, and where the isotropic Boussinesq's assumption may not be entirely valid.

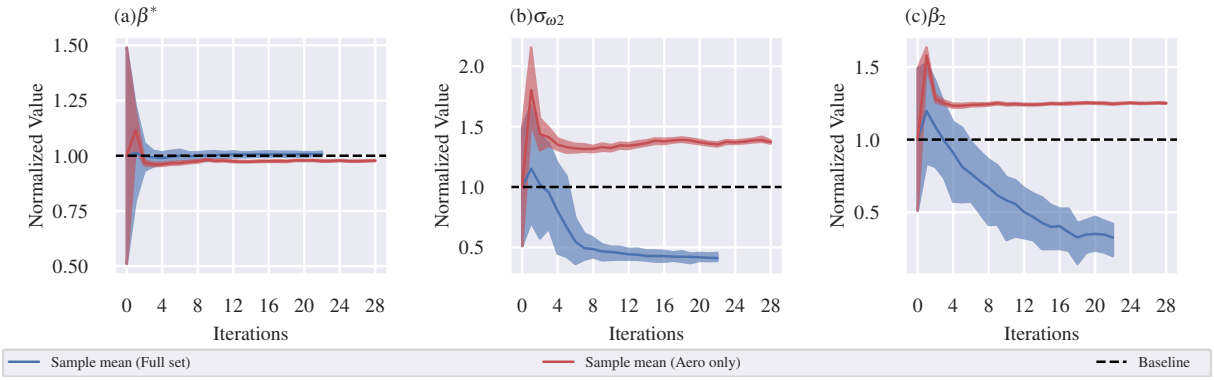
## 4. CONCLUSION

In this study, a DA technique based on an ensemble method was numerically investigated for mean density reconstruction of aero-thermal flows. In the process, 2D maps of light deviation angles obtained by synthetic BOS were considered as input observations and a RANS flow model was chosen to provide physically-plausible flow constraints. To assess the relevance of such an approach, this approximate Bayesian inference method was used as a simple model calibration tool. To validate the developed framework, a $k-\omega$ SST RANS model was considered on the application case of a hot jet impinging on a flat plate. Only a few number of coefficients in this RANS model were selected as uncertain parameters. First, two numerical twin experiments were conducted using the same RANS model, demonstrating the capability of the approach to retrieve accurate density fields and pseudo-true values of the considered parameters when model-form uncertainty is negligible. Second, results obtained with more realistic observation data issued from a LES on the same impinging jet configuration provided relatively satisfactory estimates with potential for significant improvements. These last synthetic results highlighted the importance of model-form inadequacy since the present approach can only find approximate solutions in the span of the generated prior ensemble. Despite this limitation, this iterative calibration approach is still found to be valuable when 2D BOS observations are experimentally available since it allows to reduce data mismatch after just a few number of RANS iterations compared to a much more expensive LES. In future works, this efficient approximate Bayesian estimation technique will be applied to real experimental results. Furthermore, high-dimensional parametrization will be considered to rely on more complex turbulence models (e.g. second order turbulence models such as Reynolds Stress Models). Furthermore, the inclusion of model-form uncertainties will also be investigated to provide a diagnostic tool for turbulence closure models.

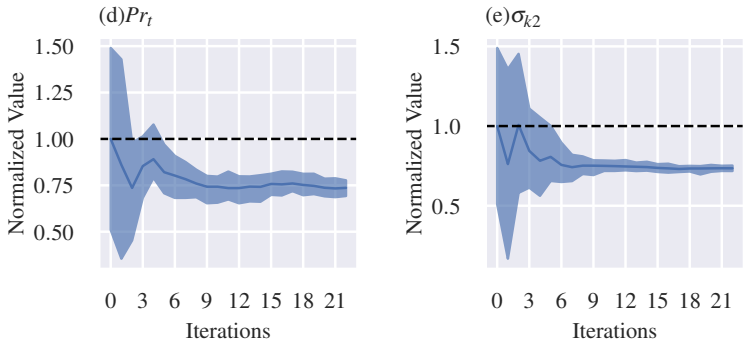
## REFERENCES

[1] B. Anderson and J. Moore. Optimal filtering. *Prentice-Hall, Englewood Cliffs, New Jersey*, 1979.

[2] C.H. Bishop et al. Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical




(A) Parameters $(\beta^*, \sigma_{\omega 2}, \beta_2)$.




(B) Parameters $(\mathrm{Pr}_t, \sigma_{k2})$ when full-set of $\Omega$ is considered.

Figure 9: Convergence of the model parameters during the assimilation procedure on the LES data for the two sets of chosen parameters, showing results for the sample and for the sample-mean.

aspects. *Monthly weather review*, 129(3):420–436, 2001.

[3] N.K. Chada et al. Parameterizations for ensemble kalman inversion. *Inverse Problems*, 34(5):055009, 2018.

[4] S.B. Dalziel et al. Whole-field density measurements by 'synthetic schlieren'. *Experiments in fluids*, 28(4):322–335, 2000.

[5] G. Evensen. *Data assimilation: the ensemble Kalman filter*, volume 2. 2009.

[6] P. Grenson. *Caractérisation expérimentale et simulations numériques d'un jet chaud impactant*. PhD thesis, Toulouse, ISAE, 2016.

[7] P. Grenson et al. Investigation of an impinging heated jet for a small nozzle-to-plate distance and high reynolds number: An extensive experimental approach. *International Journal of Heat and Mass Transfer*, 102:801–815, 2016.

[8] M.A. Iglesias et al. Ensemble kalman methods for inverse problems. *Inverse Problems*, 29(4):045001, 2013.

[9] R.E. Kalman. A new approach to linear filtering and prediction problems. 1960.

[10] M.C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[11] F.R. Menter et al. Ten years of industrial experience with the SST turbulence model. *Turbulence, heat and mass transfer*, 4(1):625–632, 2003.

[12] L. Nerger et al. PDAF-the parallel data assimilation framework: experiences with kalman filtering. In *Use of high performance computing in meteorology*, pages 63–83. 2005.

[13] F. Nicolas et al. A direct approach for instantaneous 3D density field reconstruction from background-oriented schlieren (bos) measurements. *Experiments in fluids*, 57(1):1–21, 2016.

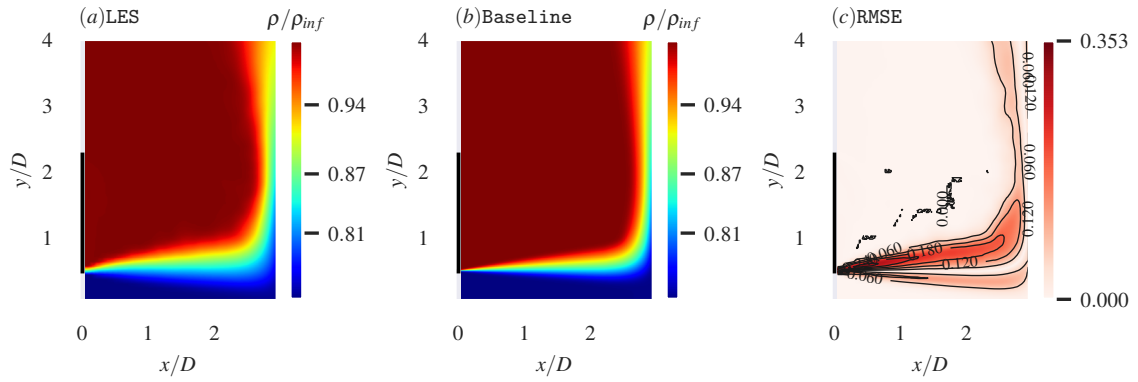



Figure 10: Density distributions of (a) the LES solution and of (b) the baseline (also the prior sample-mean). (c) Map of RMSE($\rho_{\text{LES}}, \rho_{\text{K=0}}$) between the two previous fields.

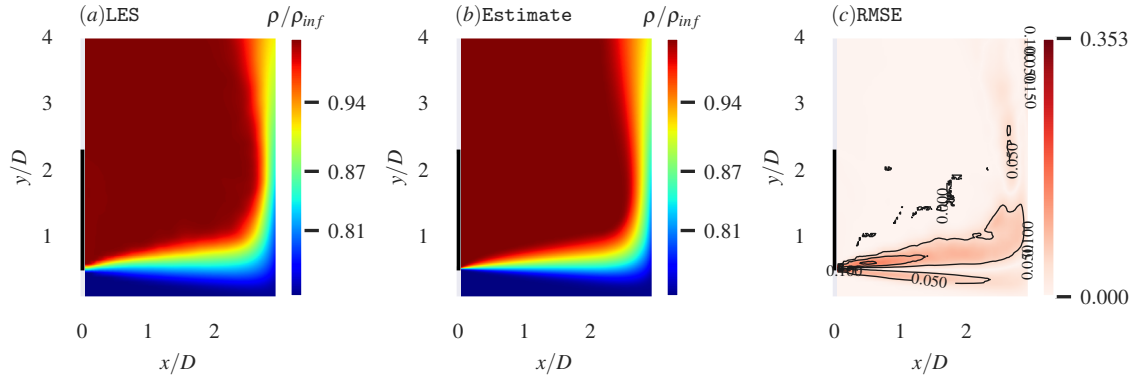



Figure 11: Density distributions of (a) the LES solution and of (b) the posterior solution (sample-mean at iteration 22). (c) Map of RMSE($\rho_{\text{LES}}, \rho_{\text{K=22}}$) between the two previous fields.

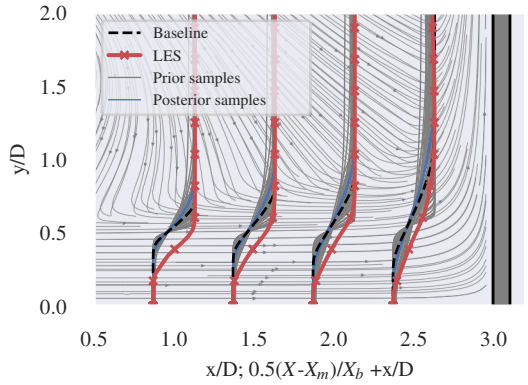



Figure 12: Density profiles across the jet at downstream locations $x/D = 1.0, 1.5, 2.0, 2.5$.

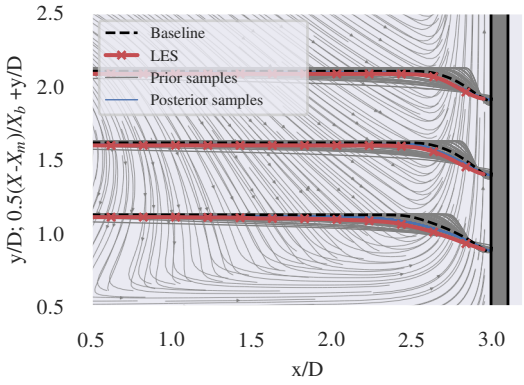



Figure 13: Density profiles across the wall jet at downstream locations $y/D = 1.0, 1.5, 2.0$.

[14] A. Refloch et al. CEDRE software. *Aerospace Lab*, (2):p–1, 2011.

[15] D.C. Wilcox et al. *Turbulence modeling for CFD*, volume 2. DCW industries La Canada, CA, 1998.

[16] H. Xiao et al. Quantifying and reducing model-form uncertainties in reynolds-averaged navier–stokes simulations: A data-driven, physics-informed bayesian approach. *Journal of Computational Physics*, 324:115–136, 2016.