Supplementary Information for

Wing morphing allows gulls to modulate static pitch stability during gliding

C. Harvey, V.B. Baliga, P. Lavoie, D.L. Altshuler

D.L. Altshuler
Email: doug@zoology.ubc.ca

This PDF file includes:

Supplementary methods
Figs. S1 to S8
Tables S1 to S3
References for SI reference citations

Supplementary methods

Model selection for elbow angle prediction. We combined our geometric morphometric analyses with machine learning techniques [1] to infer elbow and manus angles used by freely gliding gulls. We used principal components scores from the cadaver manipulations as "training" data, because each frame of the cadaver videos had both peripheral landmarks and known elbow and manus angle. We fit a variety of models (linear, general additive LOESS, random forests) and also varied the number of principal components (from two to ten) as inputs to the model to determine the relationship between wing shape and elbow or manus angle. During model training, we used a 10-fold cross-validation to inform the selection of parameters within a given model; parameters were adjusted to minimize cross-validation error. We then used two "test" data sets: (I) the set of all wind tunnel wings, and (II) wind tunnel wings with intermediate/high (>80°) elbow angles.

We found that a random forests model using scores from the first four principal components performed the best (model ID: t) (figure S2*a* and table S1). This model minimized the combined root mean square error of known elbow angles of the cadaver specimens (training data; RMS error: 1.82°), all wind tunnel wings (test set I; RMS error: 9.42°) and as well as the error of intermediate/high (elbow angle > 80°) wind tunnel wings (test set II; RMS error: 11.65°) (figure S2*a*).

Model sensitivity analysis. We performed a sensitivity analysis on our model selection which revealed that the predicted elbow range for all models is relatively constant. (figure S2*b*). Other than two models (f and k, whose RMS error was very high) all of the models predicted elbow angles that spanned the range where we measured the aerodynamic trade-off in efficiency and

stability. Therefore, our inference of elbow angles from wing shapes is not dependent on model selection.

We also investigated whether model selection would affect the result that elbow angle reduced as wind speed and wind gust increases. We found that all models (except for f and k) predict a similar relationship between elbow angle and wind speed/wind gust and have overlapping confidence intervals (model ID: t - wind speed: -0.61°/(m/s), t_{180} = -3.415, p = 0.000789; wind gust: -0.40°/(m/s), t_{180} = -3.253; p = 0.00136, figure S2*c*). It was not feasible to mark individual birds during our observational study and it is possible that individual observations are from the same bird. We also investigated the quantile regressions (evaluated at 10% intervals) of elbow angle with wind speed and wind gust respectively for our selected model (model ID: t, figure S8). For each case the 95% confidence intervals for the slope overlapped and as such were not significantly different.

Model prediction error. We investigated the prediction error of the random forests using the wings tested in the wind tunnel as a test set (since the elbow angle is known). We found the prediction error to be below 23° for the elbow and below 12° for the manus (figure S3*a, b*). This error was determined three times per wing, each time changing the plane that was parallel to the camera lens. Specifically, the: a) "full plane" made by wing tip and root chord, b) "hand plane" made of the wing tip and a chord approximately through the hand wing and, c) "interior plane" made by the root chord and the elbow joint. As these effects also tested the sensitivity of the camera and best orientation was the hand plane with a maximum error of 20°.

We explored how sensitive the prediction error was to the rotation of the wing relative to the camera using prepared wings with known elbow angles (figure S3*c, d, e, f*). The wing was rotated towards or away from the camera (positive or negative on the body axis) through a range

of approximately 100°. We found that as the wing tip was rotated about the body axis towards the camera, the error increased significantly but as the wing tip rotated in the opposite direction the error reduced. We would expect that this error would begin to increase as the wing is rotated away from the camera causing the ventral surface to be obscured from the camera. The error was less sensitive to rotations about the spanwise axis but did increase at higher rotations.

Wind tunnel specimen geometry. To calculate the wing area we determined the 3D position of 9 peripheral points (Allied Vision Technologies Prosillica GE680, 4-8 mm lens, 640 x 480 pixel resolution) and used a custom-code to rotate these points until the hand wing was flat and parallel with the horizontal and the root chord was also parallel with the horizontal (figure S4*b*). From here we used the 2D projection of these points and the bootlace method to calculate the planform area. The root chord was determined as the straight distance between points placed on the shoulder and the final secondaries (figure S4*a*). Effective wing span was determined from the 2D projection of points on the humerus and the first primary (P10) (figure S4*c*). Span-wise camber was determined from 3D imaging of 7-8 points on the leading edge of the wing. These points were rotated using rotation matrices until the vector defining the hand wing was flat and parallel to the horizontal and, because we could not track the root chord for this set up, until a point on the distal carpometacarpus was parallel with the hand wing. Results were compared to the true wings to visually assess proper rotation results (figure S4*a*). Next, to determine a comparable metric for the spanwise camber of the wings the points were rotated to ensure the humerus and the distal P10 point were flat on the horizon. From this orientation the maximum spanwise camber was determined to be the maximum height of the leading edge over the wing span (figure S1).

Experimental uncertainty analysis. We calculated the uncertainty of the time series readings of the pressure, force and moment data using the integrated autocorrelation time calculated by the R package dirmcmc [2]. This result was propagated through all calculations. Additionally, this propagation included error of wing area (4%), root chord (3%) and bias errors from machinery. Error from measurements and equipment was also included within the uncertainty propagation as applicable. Uncertainty of the results are displayed as error bars in figure 3. The uncertainty is minimal for all results except for the maximum coefficient of lift.

Estimating induced drag over wing root. The induced drag in our experiment is expected to be overestimated due to the gap between the root of the wings and the tunnel walls (figure S5). We estimated this effect by using the standard planar wing equation to compute the amount of induced drag that would occur if that the half model was truly a full model (using a span efficiency of 1)[3]. In this simplified case we find that the induced drag over the root is equivalent to the tip effect. We removed that value from the measured coefficient of drag and re-computed the aerodynamic efficiency (figure S5*a*). We found that the relative difference between our estimation and measured value was 5-20% (figure S5*b*). Wings were tested at varying distances from the wall to allow us to ensure that the hand wing remained approximately perpendicular to the freestream flow (figure S5*c*). This method provides a conservative estimate because the presence of the wall in proximity to the root will have a damping effect compared to the wing in an infinite environment, which is assumed in our model. Induced drag over the wing root may be further reduced due to the geometry of the wings, as span-wise camber improves the span efficiency above 1 and reduces the effects of induced drag compared to their planar counterparts [4]. Finally, it is possible that the porous feather structure may allow air to pass through across the wing to interact with the boundary layer and impact the induced drag [5]. As

a result of these findings, we expect that our results will hold despite the potential over-estimation of drag. Induced drag has a negligible effect on the pitching moment due to its generally small magnitude of the force and small moment arm. Thus an increase of the induced drag in itself should not substantially affect the pitching moment measured in our experiment.

Aerodynamic results statistical analysis. The pitch stability derivative (figure 3*f*) was determined by fitting a linear model to the linear range of data pre-stall and computing the slope of the line. 95% confidence intervals were computed and displayed.

We used linear mixed-effects models from R package nlme [6] that allowed us to account for repeated measures on the same wings over different turbulence intensities. To do this we first selected the models by comparing different linear, quadratic, and cubic fits using the Akaike information criterion and Akaike weights (table S2). We found the best fit was a linear exponent on elbow angle for the maximum coefficient of lift and coefficient of pitch at the aerodynamic center. We found that the quadratic fit best explained the minimum coefficient of drag and the cubic model best predicted the aerodynamic efficiency and pitch stability derivative.

We next used an ANOVA test to compare a reduced model (signifying the null hypothesis) to the full model to determine the association between our explanatory variables (elbow angle and turbulence intensity) for each response variable (aerodynamic characteristics) (table S3). All comparisons (except for the effect of turbulence intensity on the minimum drag) allowed us to reject the null hypothesis and find that there is a significant probability that there is an association between our tested explanatory and response variables. The medians of these models were visualized for figure 3*b, c, e* and *g* using a model-based parametric bootstrap method with 100 simulations from the lme4 package [7].

Turbulence Grids/Calculations. To examine effects of turbulence intensity on gull wings, we varied turbulence intensity using turbulence generating grids. We had three conditions: a) low turbulence, no grid b) low/medium turbulence, grid ID: Rd38 c) medium turbulence, grid ID: Sq39. Rd38 is a bi-planar round rod grid consisting of round aluminum rods with a 6.8 mm nominal diameter mounted in a machined frame with a mesh length of 32 mm. Sq39 is a square mesh consisting of a single piece of 6.35 mm thick aluminum that was water cut with a mesh length of 100 mm. The two grids generate a uniform mean flow with homogeneous turbulence with integral length scales of approximately 16 mm for Rd38 and 43 mm for Sq39 as verified in previous studies [8].

For each turbulence grid, we used a single hot-wire probe operated with a constant temperature anemometer built by the University of Newcastle [9] to measure turbulence intensity at the stream-wise plane approximately at the leading edge of the mounted wings. The hot-wire was calibrated against 19 known velocities spanning 2 m/s and 20 m/s, to which Kings Law was fitted. The calibration points were sampled at 20kHz for 120 seconds. The turbulence intensity was calculated as the root mean square of the velocity fluctuations over the average velocity. The one-sided power spectral density of the velocity time series was determined from the fast Fourier transform of the signal. These results were normalized to the streamwise velocity and the mean root chord of the models tested, which is approximately 0.21 m (figure S7*a*). The total mean-square fluctuation of the time series is given by the integral of the spectral density (figure S7*b*). The calculated values agreed with results from previous studies [8]. We found that the energy peaks (figure S7*b*) were close to the convective time scale of the wings.

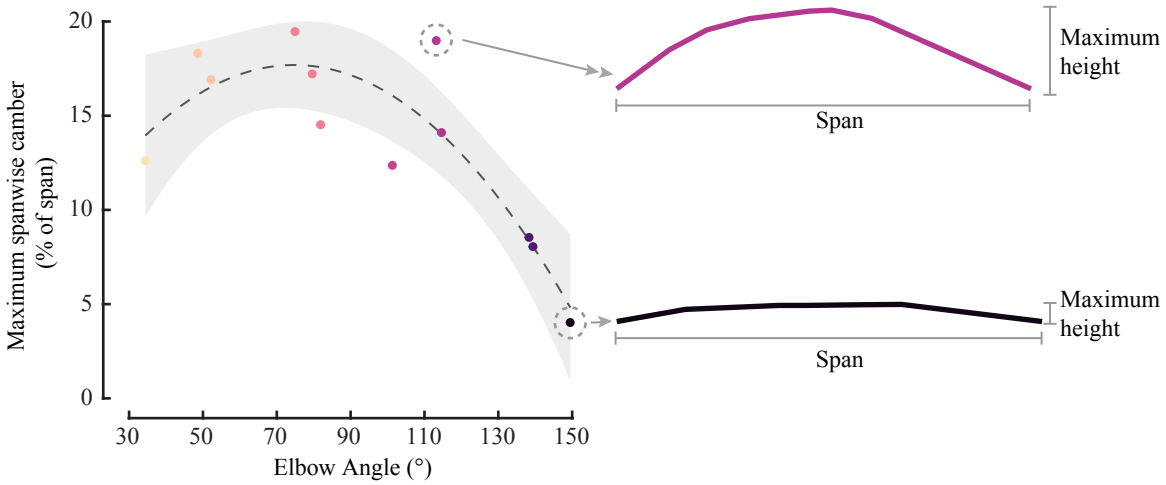

Figure S1. Spanwise camber increases with elbow angle across the *in vivo* range. Maximum spanwise camber is determined as the ratio of maximum height of the leading edge to the wing span.

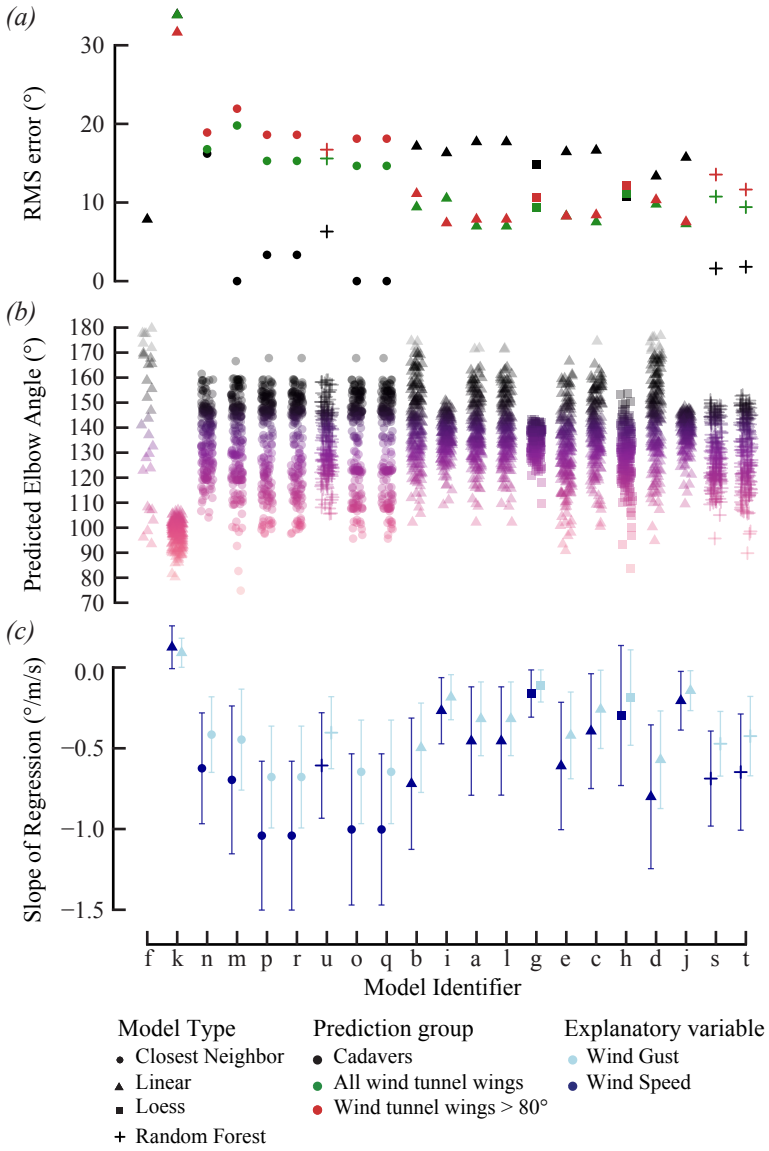

Figure S2. Sensitivity analysis of the prediction model selection We investigated the sensitivity of our results to the selected model used to predict elbow angle. We compared 23 models including linear, loess, random forests and an estimator based on the closest neighbor (Table S1). (*a*) Root-mean-square (RMS) error of the elbow angle prediction was computed for three groups: a) stills from cadaver manipulations b) photos of wind tunnel wings c) a subset of the wind tunnel wings above 80°. We then scored the models using the RMS error from the three plotted errors and this informed the ordering of the x axis from worst to best. We selected to use the model with the lowest RMS error (random forests model with four principal components). (*b*) Predictions of in vivo elbow angles from each model. The predicted range for all models (except for *f* and *k*) spans our range of aerodynamic trade-offs and choice of model does not affect our conclusions. (*c*) The predicted elbow angle also indicates that there is a trend with wind speed and wind gust. Again we found that despite the selected model we would arrive at approximately the same results (except for *f* and *k*).

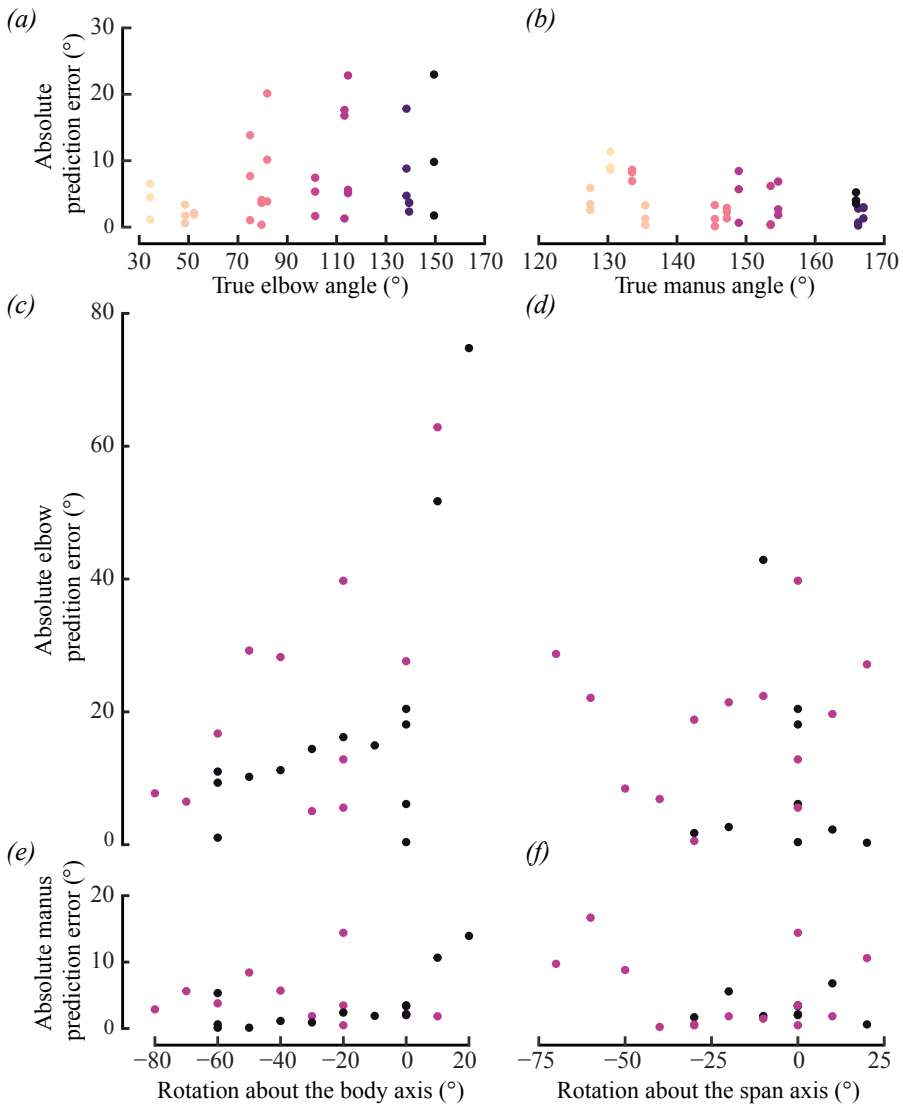

Figure S3. Error of the selected prediction model and sensitivity analysis of camera position relative to the wing. A random forests model using the first four principal components was fit to results from the cadaver manipulation and error was assessed by using the model to predict wings with known elbow and manus angles. (*a*) Elbow angle prediction was assessed using three different camera perspectives and we found 23° maximum absolute error and 6.9° average error. (*b*) Manus angle prediction had 12° maximum absolute error and 3.8° average error. Sensitivity to the perpendicularity was investigated and, when predicting elbow angles, (*c*) found that error grew as the wing rotated about the body axis and was relatively constant as the wing rotated away from the camera about the same axis. (*d*) The model was less sensitive to rotations about the span-axis. (*e*, *f*) Similar yet diminished trends in sensitivity to perpendicularity were observed for manus angle predictions with maximum absolute error always remaining under 20°. The model had errors over 40° only for photos that were clearly not a ventral view. The sensitivity analysis was completed with wings of known elbow angles of 149° (black dots) and 108° (purple dots).

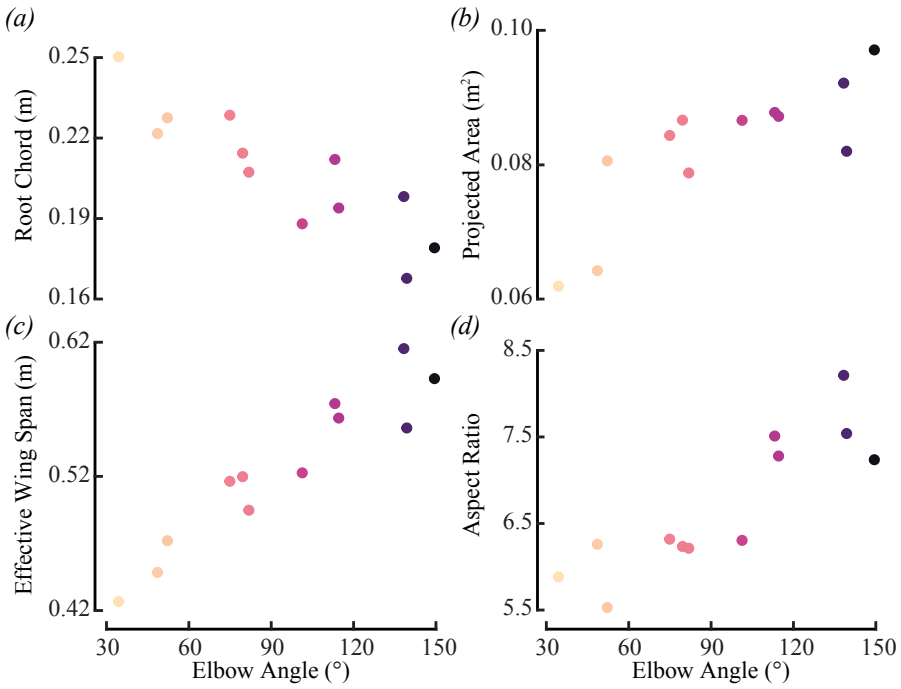

Figure S4. The geometric properties of the wind tunnel wing specimens. (*a*) Chord measured at the root, (*b*) 2D projected area of the wings, (*c*) effective wing span, (*d*) aspect ratio based on effective area and effective span ($AR = eff\ wing\ span^2/(2 \cdot eff\ wing\ area)$).

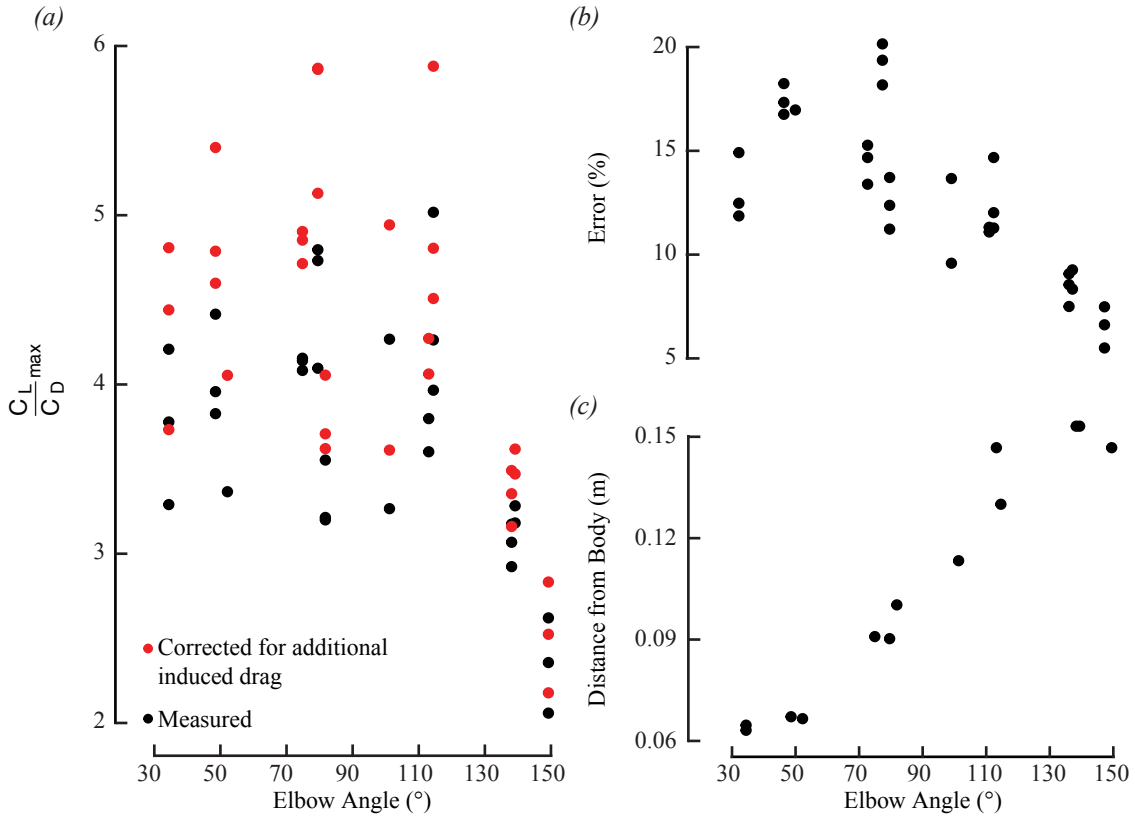

Figure S5. Effects of induced drag over the wing root. Size constraints of the wind tunnel allowed only a single wing to be tested at a time. This set-up potentially allowed for additional induced drag to be incurred due to airflow around the root of the wing. (*a*) This effect is estimated using the standard fixed wing induced drag calculation with a span efficiency of 1. (*b*) This analysis revealed that 5-20% relative error in the aerodynamic efficiency was possible. (*c*) The wings were tested at varying distances from the wall to allow correct positioning of the leading edge. The wall would have a dampening effect on the induced drag and lessen the error.

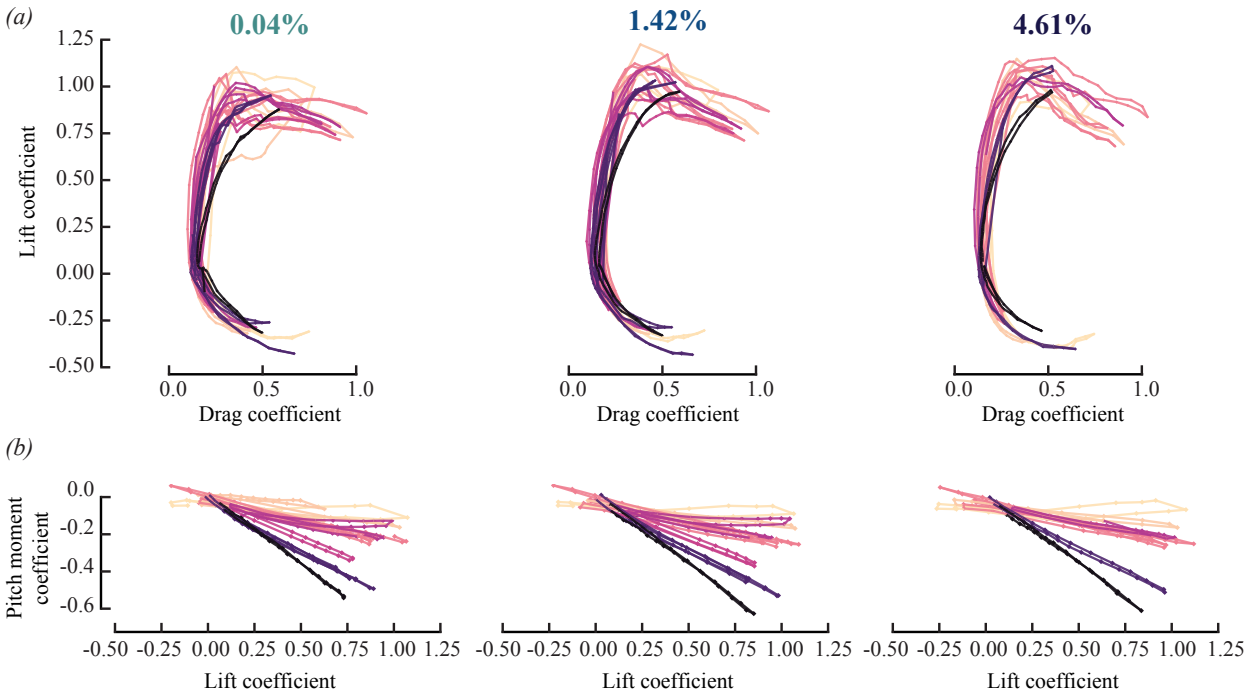

Figure S6. Wind tunnel results for all wings tested over the three turbulence intensities. (*a*) Lift-to-drag polars (*b*) Pitching moment coefficient variation with coefficient of lift.

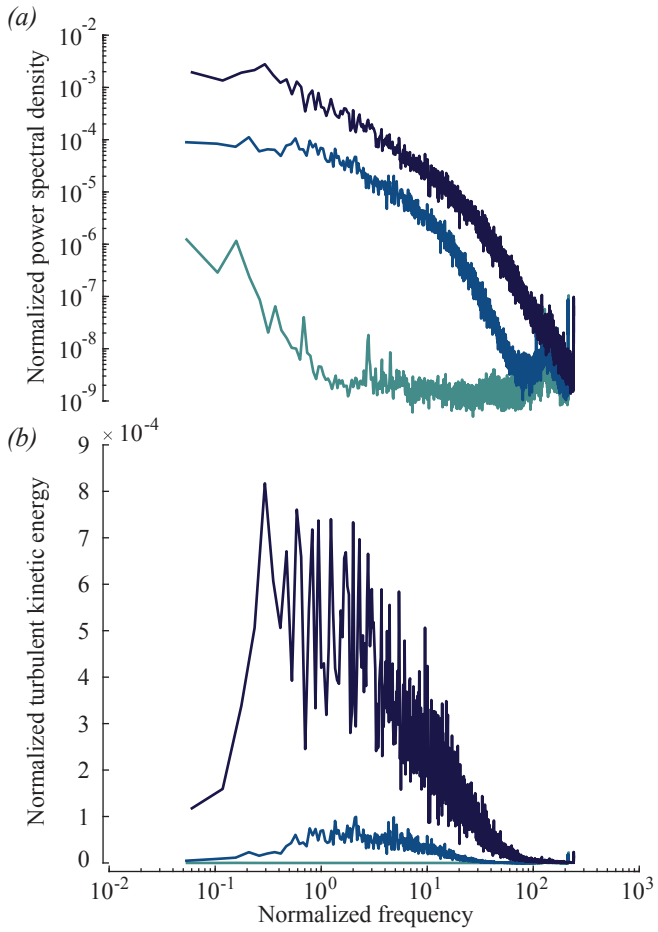

Fig. S7. Turbulence intensity was measured through hot-wire. (*a*) Turbulence intensity was calculated from the power spectral density as the root-mean-square of velocity by the mean streamwise velocity. Freestream turbulence intensity with no grid was 0.04%, the medium grid had 1.42% turbulence intensity and the coarse grid high turbulence intensity 4.61%. (*b*) The compensated semi-log plot is a visually proportional representation of the true distribution of the turbulent kinetic energy over the relevant time scales.

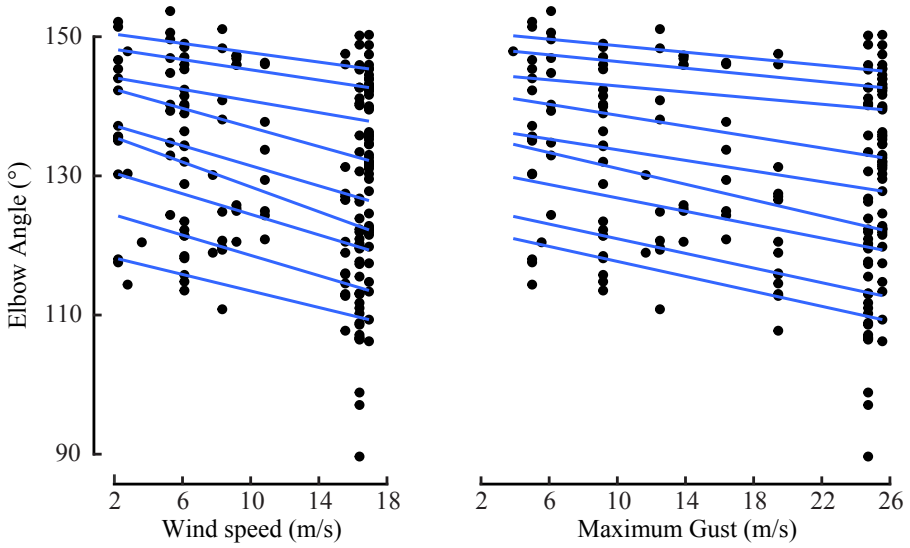

Fig. S8. Quantile regression of the wind data vs. observed elbow angle. The regression coefficients of the quantiles evaluated at 10% intervals do not differ significantly.

Table S1. Sensitivity analysis of the model selection.

ID	Model equation	RMS error (°)				Slope (°/m/s)	
		cadaver	*Test I*	*Test II*	*score*	*Wind*	*Gust*
a	lm(elbow.angle ~ PC1 + PC2)	17.72	6.99	7.86	20.60	-0.46	-0.32
b	lm(elbow.angle ~ PC1 * PC2)	17.15	9.42	11.13	22.51	-0.72	-0.50
c	lm(elbow.angle ~ PC1 + PC2 + PC3 +PC4)	16.63	7.52	8.42	20.10	-0.39	-0.26
d	lm(elbow.angle ~ PC1 * PC2 * PC3 *PC4)	13.35	9.78	10.34	19.51	-0.80	-0.57
e	lm(elbow.angle ~ PC1 + PC2 + PC3 + PC4 + PC5 + PC6 + PC7 + PC8 + PC9 + PC10)	16.45	8.28	8.26	20.18	-0.61	-0.42
f	lm(elbow.angle ~ PC1 * PC2 * PC3 * PC4 * PC5 * PC6 * PC7 * PC8 * PC9 *PC10)	7.86	749.39	849.55	1132.86	-5964.79	-4047.11
g	loess(elbow.angle ~ PC1 + PC2)	14.82	9.35	10.70	20.53	-0.16	-0.11
h	loess(elbow.angle ~ PC1 + PC2 + PC3 +PC4)	10.75	11.12	12.18	19.69	-0.30	-0.19
i	lm(elbow.angle ~ PC1^2 + PC2^2)	16.31	10.53	7.40	20.78	-0.27	-0.18
j	lm(elbow.angle ~ PC1^3 + PC2^3)	15.73	7.28	7.53	18.90	-0.21	-0.14
k	lm(elbow.angle ~ cos(PC1) + cos(PC2))	33.88	33.85	31.65	57.41	0.13	0.09
l	lm(elbow.angle ~ sin(PC1) + sin(PC2))	17.71	6.99	7.87	20.60	-0.46	-0.32
m	Closest Neighbor - 2PCs	0.00	19.79	21.93	29.54	-0.70	-0.45
n	Closest Neighbor - 2PCs (Mean of 3)	16.22	16.78	18.90	30.03	-0.62	-0.42
o	Closest Neighbor - 6PCs	0.00	14.66	18.12	23.31	-1.00	-0.65
p	Closest Neighbor - 6PCs (Mean of 3)	3.34	15.29	18.61	24.32	-1.04	-0.68
q	Closest Neighbor - 10PCs	0.00	14.66	18.12	23.31	-1.00	-0.65
r	Closest Neighbor - 10PCs (Mean of 3)	3.34	15.29	18.61	24.32	-1.04	-0.68
s	rf(elbow.angle ~ PC1 + PC2)	1.61	10.75	13.55	17.38	-0.69	-0.47
t	rf(elbow.angle ~ PC1 + PC2 + PC3 +PC4)	1.82	9.42	11.65	15.09	-0.65	-0.42
u	rf(elbow.angle ~ PC1 + PC2 + PC3 + PC4 + PC5 + PC6 + PC7 + PC8 + PC9 + PC10)	6.31	15.60	16.72	23.72	-0.61	-0.40

We tested 23 models and predicted elbow angles from fully characterized wings. We used the root-mean-square error from the prediction groups to define an overall score for each model. The four PC random forest model (gray shading) best minimized the error between the three categories. Additionally, we computed the regression coefficient for how elbow angle changed with wind and gust speed and found minimal differences between models (figure S2).

Table S2. Model selection for aerodynamic parameters

		~ ElbowAngle^_ + TI + (1|WingID)		
		1	2	3
C_{Lmax}	*AIC*	-74.96	-73.05	-71.05
	AIC differences	0.00	1.91	3.91
	Relative likelihood of model	1.00	0.38	0.14
	Akaike weights	0.66	0.25	0.09
$(L/D)_{max}$	*AIC*	47.71	42.67	41.92
	AIC differences	5.79	0.76	0.00
	Relative likelihood of model	0.06	0.69	1.00
	Akaike weights	0.03	0.39	0.57
C_{Dmin}	*AIC*	-157.44	-162.99	-161.76
	AIC differences	5.55	0.00	1.23
	Relative likelihood of model	0.06	1.00	0.54
	Akaike weights	0.04	0.62	0.34
dC_m/dC_L	*AIC*	-124.90	-124.05	-125.90
	AIC differences	0.99	1.84	0.00
	Relative likelihood of model	0.61	0.40	1.00
	Akaike weights	0.30	0.20	0.50
C_{mac}	*AIC*	-171.71	-169.73	-170.03
	AIC differences	0.00	1.98	1.68
	Relative likelihood of model	1.00	0.37	0.43
	Akaike weights	0.55	0.21	0.24

Using the Akaike information criterion (AIC) and Akaike weights we investigated the best linear polynomial fit. Columns represent the different exponents on elbow angle for each model tested.

Table S3. ANOVA comparison between full and reduced model.

Model	$\chi 2$	Df	Pr(>$\chi 2$)	
Elbow angle				
Aerodynamic Efficiency ~ 1 + TI + (1 | WingID)	13.48	3	0.0037	**
Aerodynamic Efficiency ~ ElbowAngle^3 + TI + (1 | WingID)				
Maximum lift coefficient ~ 1 + TI + (1 | WingID)	5.003	1	0.0253	*
Maximum lift coefficient ~ ElbowAngle + TI + (1 | WingID)				
Minimum drag coefficient ~ 1 + TI + (1 | WingID)	10.29	2	0.0058	**
Minimum drag coefficient ~ ElbowAngle^2 + TI + (1 | WingID)				
Pitch stability derivative ~ 1 + TI + (1 | WingID)	18.48	3	0.0003	***
Pitch stability derivative ~ ElbowAngle^3 + TI + (1 | WingID)				
Zero-lift pitch ~ 1 + TI + (1 | WingID)	4.39	1	0.0362	*
Zero-lift pitch ~ ElbowAngle + TI + (1 | WingID)				
Turbulence intensity				
Aerodynamic Efficiency ~ ElbowAngle^3 + 1 + (1 | WingID)	5.942	1	0.0147	*
Aerodynamic Efficiency ~ ElbowAngle^3 + TI + (1 | WingID)				
Maximum lift coefficient ~ ElbowAngle + 1 + (1 | WingID)	12.9	1	0.0003	***
Maximum lift coefficient ~ ElbowAngle + TI + (1 | WingID)				
Minimum drag coefficient ~ ElbowAngle^2 + 1 + (1 | WingID)	0.172	1	0.6781	
Minimum drag coefficient ~ ElbowAngle^2 + TI + (1 | WingID)				
Pitch stability derivative ~ ElbowAngle^3 + 1 + (1 | WingID)	34.03	1	<0.0001	***
Pitch stability derivative ~ ElbowAngle^3 + TI + (1 | WingID)				
Zero-lift pitch ~ ElbowAngle + 1 + (1 | WingID)	18.68	1	<0.0001	***
Zero-lift pitch ~ ElbowAngle + TI + (1 | WingID)				

Using the model selected by AIC (table S2) we used ANOVA to compare the linear mixed-effect models to a reduced model (gray font) to determine the statistical significance of the explanatory variable (either elbow angle or turbulence intensity). A p-value of less than 0.05 indicates that the explanatory variable has a significant effect on the response variable.

References

1. James G, Witten D, Hastie T. 2013 *An introduction to statistical learning*. Springer-Verlag New York.

2. Mallik A. 2017 *Directional metropolis hastings algorithm*. R package,.

3. Anderson Jr JD. 1989 *Introduction to Flight*. McGraw-Hill Higher Education.

4. Lazos B, Visser K. 2006 Aerodynamic comparison of hyper-elliptic cambered span (HECS) wings with conventional configurations. In *24th AIAA Applied Aerodynamics Conference*, pp. 1–18. American Institute of Aeronautics and Astronautics.

5. Eder H, Fiedler W, Pascoe X. 2011 Air-permeable hole-pattern and nose-droop control improve aerodynamic performance of primary feathers. *J. Comp. Physiol. A* **197**, 109–117.

6. Pinheiro J, Bates D, DebRoy S, Sarkar D, {R Core Team}. 2017 *{nlme}: Linear and nonlinear mixed effects models*.

7. Bates D, Maechler M, Bolker B, Walker S. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48. (doi:10.18637/jss.v067.i01)

8. Hearst RJ, Lavoie P. 2016 Effects of multi-scale and regular grid geometries on decaying turbulence. *J. Fluid Mech.* **803**, 528–555.

9. Miller IS, Shah DA, Antonia RA. 1987 A constant temperature hot-wire anemometer. *J. Phys. [E]* **20**, 311.